Measurement of W^+W^- production cross section in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

DISSERTATION

zur Erlangung des akademischen Grades

DOCTOR RERUM NATURALIUM

(Dr. rer. nat.) im Fach Physik

Spezialisierung: Experimentalphysik

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von

M.Sc. Baishali Dutta

Präsidentin der Humboldt-Universität zu Berlin
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät
Prof. Dr. Elmar Kulke

Gutachter/innen:
1. PD Dr. Klaus Mönig
2. Prof. Dr. Thomas Lohse
3. PD Dr. Matthias Mozer

Tag der mündlichen Prüfung: 19.06.2018

Acknowledgements

First of all, I would like to express my sincere gratitude to my supervisor Prof. Klaus Mönig for the support of my research work, motivation, invaluable advice and especially for believing in me. His guidance helped me a lot in flourishing my understanding of Elementary Particle Physics and also pursuing my doctorate.

Besides my supervisor, I would like to immensely thank my advisor Dr. Kristin Lohwasser for her insightful comments and encouragement in my research work. She not only helped me in my research but also for the hard questions which incented me to widen my research from various perspectives. Moreover, I am incredibly grateful to her for having faith in me whenever I was troubled. I could not have imagined having a better advisor and mentor for my research work.

I cordially thank my colleagues, especially Dr. Thorsten Kuhl and fellow PhD students Akanksha and Elias for the stimulating discussions and for all the fun we have had during my PhD period at DESY Zeuthen. Also, I thank all my friends for their continuous moral boost and for being there in my ups and downs.

Above all, I would like to thank my family - my parents, my elder sister and my best friend Amit for supporting me spiritually throughout my journey as a doctoral student and my life in general.

Zusammenfassung

Diese Arbeit stellt eine Messung des W^+W^- Produktionswirkungsquerschnitts in pp Kollisionen bei einer Schwerpunktenergie von $\sqrt{s} = 13$ TeV vor. Der verwendete Datensatz wurde mit dem ATLAS Detektor im Jahr 2015 aufgezeichnet und entspricht einer integrierten Luminosität von 3,16 fb^{-1}. Die W^+W^- Ereignisse werden über leptonische Zerfälle der W-Bosonen selektiert, wobei jeweils eines der W Bosonen in ein Elektron und ein Elektronneutrino und das andere in ein Myon und ein Myonneutrino zerfällt. Der gemessene Wirkungsquerschnitt in einem Referenzphasenraum, der nahe der Detektorakzeptanz definiert ist, beträgt

$$\sigma_{W^+W^-}^{\text{fiducial}} = 529 \pm 20\,(\text{stat.}) \pm 50\,(\text{syst.}) \pm 11\,(\text{lumi.})\,\text{fb}.$$

Das Ergebnis stimmt innerhalb der Fehlergrenzen mit der besten verfügbaren Standardmodell-Vorhersage von 478 ± 17 fb überein. Die kinematischen Verteilungen der im Zerfall der beiden W Bosonen produzierten Leptonen wurden genauer untersucht, um die drei-Eichbosonenkopplung bei den $WW\gamma$ und WWZ Vertizes zu messen. Abweichungen dieser Kopplungen vom Standardmodell können ein Hinweis auf neue Physik sein. Für die Größen, welche diese anomalen Kopplungen parametrisieren, wurden Konfidenzintervalle berechnet. Die beobachteten Kopplungen stimmen mit dem Standardmodell überein.

Abstract

This thesis presents a measurement of W^+W^- production cross section with the pp collision data collected at the ATLAS detector in the year of 2015. The dataset corresponds to a centre-of-mass collision energy of $\sqrt{s} = 13$ TeV with a total integrated luminosity of 3.16 fb^{-1}. The W^+W^- signal events are selected using a signature where one of the W bosons decays into an electron and an electron neutrino while the other produces a muon with an associated muon neutrino. The measured cross section in the chosen fiducial phase space close to detector acceptance is

$$\sigma^{\text{fiducial}}_{W^+W^-} = 529 \pm 20\,(\text{stat.}) \pm 50\,(\text{syst.}) \pm 11\,(\text{lumi.})\,\text{fb}.$$

The result within the assigned uncertainties is compatible with the best available Standard Model prediction of 478 ± 17 fb. The observed kinematic spectrums of the produced leptons from the decay of the two W bosons are further investigated to study the triple gauge boson couplings at the $WW\gamma$ and WWZ vertices. The deviation of these couplings from the Standard Model can probe the existence of new physics. The confidence intervals have been calculated for the parameters representing these anomalous couplings. The observations are consistent with the Standard Model expectations.

Contents

Chapter 1.

Introduction

A revolutionary era for the particle physics experiments at the energy frontier commenced in March 2010 with the first collision at the Large Hadron Collider (LHC). The LHC with its unprecedented energy reach accomplished one of its primary goals, the discovery of the Higgs boson predicted in the Standard Model (SM) within just over two years. In July 2012, the two main LHC experiments ATLAS and CMS announced the observation of this SM-like Higgs boson with five sigma signifcance [1, 2]. The high energy proton-proton collisions at the LHC further provide prospects to study the properties of the discovered SM Higgs boson, understand the SM more precisely as well as search for new physics beyond the SM.

The SM is known to be a well-tested theoretical model with no contravening experimental observations so far. The validity of the SM is proven till date at the currently attainable accelerator energies. However, there exist many theoretical questions unexplained by the SM that imply the incompleteness of this theory. Such questions address the baryon asymmetry of the universe, the presence of the dark matter and dark energy and the masses of the neutrinos. Moreover, the SM Higgs boson mass of $\sim$125 GeV [3] is very small compared to the Planck mass, referred to as the "hierarchy problem", is yet to be explained. These suggest the existence of unobserved new physics phenomena. Since the nature of this new physics is not known, it becomes necessary to test the SM with higher precisions at high energy physics experiments and put more stringent constraints on its parameters highly sensitive to deviations from the SM expectations.

A precision measurement of the W^+W^- (denoted as WW for simplicity) production cross section offers a unique platform to test the SM predictions while the anomalous triple gauge boson couplings at the $WW\gamma$ and WWZ vertices further constraint the SM parameter space and probe new physics.

The WW production cross section measurements at the ATLAS and CMS experiments at the centre-of-mass energies of 7 and 8 TeV [4–7] have induced a large interest amongst the particle physics enthusiasts in the recent few years. These measurements consistently observed an excess of the WW cross section in comparison with the next-to-leading order (NLO) SM prediction in the perturbative quantum chromodynamics. The discrepancies between the measurements

and predictions are reduced with higher order theoretical calculations at next-to-next-to-leading order (NNLO) and also taking into account the logarithmic contributions resulting from soft gluon emission at next-to-next-to-leading logarithm (NNLL) accuracy [8–11]. Therefore, it would be fascinating to confront the best available theoretical prediction with an updated measurement at a collision energy of 13 TeV, almost twice as large as before.

Besides, the WW production process contributes as a major background source to the Higgs boson decaying into the WW final state. It is therefore important to improve our understanding and modelling of the WW process for a better understanding of the Higgs boson properties.

The production of the W boson pair provides a crucial test for the electroweak sector in the SM. The gauge bosons of the electroweak interaction couple with each other due to the *non-abelian* gauge structure of the underlying electroweak theory. This characteristic permits vertices with three or four gauge boson couplings. The SM precisely predicts these couplings where some of them vanish in the SM. However, these couplings are experimentally not well constrained. The WW pair production through a virtual photon γ or Z boson allows us to study the triple gauge boson couplings at the $WW\gamma$ and WWZ vertices.

The implication of this *non-abelian* structure further influences the WW production rate at higher energies. The two dominant subprocesses of the WW production involve the interaction of the quarks - the proton constituents where the W boson pair is produced by the scattering process as well as by the annihilation process through the exchange of a virtual photon γ or Z boson. The latter involves the self-coupling of the gauge bosons as mentioned earlier. The production rates of both these subprocesses increase to infinitely large values with higher energies. However, the negative inference between these two nullifies these divergences. This is only possible if the coupling strengths of the W bosons to other particles in the SM and its self-couplings are compliant with the SM predictions. Therefore, the WW production is a challenging process enabling us to test the fundamental structure of the SM itself.

Hence, the measurement of the WW production cross section presented in this thesis provides a means to test the predictability of the SM and study the sensitivity of these triple gauge boson couplings deviating from the SM predictions.

This chapter provides an overview and motivations of the studies presented in this thesis. Chapter 2 discusses the theoretical knowledge required to study the production of the WW process in the proton-proton collisions at the LHC. Chapter 3 describes the ATLAS detector and its subsystems that collect the proton-proton collision events which are further processed through the event reconstruction procedures as elaborated in Chapter 4. The WW signal selection and contributing background estimations are described in Chapter 5, followed by the WW cross section measurement and its comparison with the SM expectation in Chapter 6. The study of the anomalous triple gauge boson couplings is detailed in Chapter 7 and finally Chapter 8 concludes with a summary of the analysis presented in this thesis work.

Chapter 2.

Theoretical Motivation

The theoretical knowledge to understand the WW production process is reviewed in this chapter.

Section 2.1 discusses the Standard Model (SM) [12–16] of particle physics, proven to be the most successful theoretical model so far to describe physics interactions in the microscopic realm. Developed mainly during 1960's and 1970's, it remarkably explains the three fundamental interactions - the electromagnetic, weak and strong interactions of nature that govern our universe. The fourth fundamental interaction, the gravitational force has so far not been successfully incorporated into the SM. However, at the typical scale of elementary particle physics, the strength of the gravitational force is far too weak even in comparison with the weak interaction (the weakest force in the SM) to play a significant role [17]. Therefore, it can be ignored for the results presented in this thesis. The particles and nature of physics interactions predicted by the SM are verified over the years at significantly high precisions in the high energy physics experiments [3]. This section provides an insight into the nature of interactions amongst the particles in the SM.

This analysis studies the WW production process using the proton-proton (pp) collision dataset provided by the LHC. The phenomenology of a pp collision and different phases of its evolution are therefore discussed in Section 2.2. The generation of Monte Carlo (MC) samples as detailed in Section 2.3 considers all these evolution phases to model physics processes that contribute to an analysis.

Finally, Section 2.4 presents the production mechanisms and important aspects of the $pp \to WW$ process whereas Section 2.5 describes the theoretical concept of the anomalous triple gauge boson couplings that can probe physics beyond the Standard Model.

2.1. The Standard Model

The Standard Model is mathematically formulated as a quantum field theory that combines the special theory of relativity and quantum mechanics. Physical matter is composed of fundamental spin-$\frac{1}{2}$ particles, called fermions. In this field theory approach, the fermions

are defined as local excitations of the underlying fields. The forces of nature occur from the interactions between these fundamental fields which are introduced in the SM as gauge fields belonging to certain symmetry groups. The quanta (excitations) of these gauge fields are the spin-1 bosons, named as the gauge bosons in the SM. The gauge bosons mediate the forces between the matter particles (fermions). The symmetry carried by an interaction is obtained by requiring the Lagrangian describing the interaction to be invariant under local (space-time) transformation of the corresponding symmetry group and is responsible for conserving specific physical quantities [18]. The combined symmetry group representing the internal symmetry of the SM is given as

$$U(1)_Y \otimes SU(2)_L \otimes SU(3)_C$$

where the electromagnetic and weak interactions in the SM are combinedly represented by the $U(1)_Y \otimes SU(2)_L$ gauge group, spontaneously broken by the Higgs mechanism as discussed in Subsection 2.1.1. The quantum field theory that illustrates the electromagnetic interaction is called Quantum Electrodynamics (QED). The strong interaction is described by the $SU(3)_C$ symmetry group and the corresponding field theory is referred to as the Quantum Chromodynamics (QCD), addressed in Subsection 2.1.2.

Elementary Particles and Forces

The elementary particles and nature of forces existing in the SM are introduced in the following. Each particle whether elementary or composite has its own antiparticle. An antiparticle exhibits identical properties but has opposite sign of its associated quantum numbers and electrical charge to that of the particle.

The fermions, basic constituents of physical matter, are classified into two types of particles - leptons and quarks. Both types are arranged in three generations where the particles in one generation differ from the particles in other generations by their masses and the properties depending on the mass, such as lifetime and kinematically allowed decay modes.

Each lepton generation includes two leptons - a particle with an electrical charge of -1 and an electrically neutral neutrino particle. The charged leptons are the electron, muon and tau particles that are associated with the corresponding electron, muon and tau neutrinos respectively.

The quarks are also split into three generations - (up, down), (charm, strange) and (top, bottom) that carry electrical charges of ($+\frac{2}{3}$, $-\frac{1}{3}$). The quarks carry a quantum number called colour charge arbitrarily chosen to be red, green or blue. In nature, only colour-neutral particles are observed. Therefore, the quarks are never witnessed in isolation rather in the form of hadrons - mesons and baryons. The mesons are composed of a quark with a certain colour charge and an antiquark with the corresponding anticolour charge. The baryons are formed as a triplet of

quarks with associated colour charges of red, green and blue which by definition sum up to a colour-neutral state.

The quarks and charged leptons are massive but the neutrinos are treated as massless particles in the SM. Although the observed neutrino oscillations [19] indicate the neutrinos to have non-vanishing masses as well. The magnitude of the three neutrino masses have not yet been measured but neutrino experiments provide upper limits [3]. Table 2.1 shows the fundamental fermions together with some of their essential properties.

Generation	Flavour	Symbol	Mass	Charge	Flavour	Symbol	Mass	Charge
	Quarks				Leptons			
1^{st}	Up	u	$2.2^{+0.6}_{-0.4}$ MeV	$+\frac{2}{3}$	Electron	e	0.51 MeV	-1
	Down	d	$4.7^{+0.5}_{-0.4}$ MeV	$-\frac{1}{3}$	Electron neutrino	ν_e	< 2 eV	0
2^{nd}	Charm	c	1.28 ± 0.03 GeV	$+\frac{2}{3}$	Muon	μ	105.66 MeV	-1
	Strange	s	96^{+8}_{-4} MeV	$-\frac{1}{3}$	Muon neutrino	ν_μ	< 2 eV	0
3^{rd}	Top	t	173.1 ± 0.6 GeV	$+\frac{2}{3}$	Tau	τ	1776.86 ± 0.12 MeV	-1
	Bottom	b	$4.18^{+0.04}_{-0.03}$ GeV	$-\frac{1}{3}$	Tau neutrino	ν_τ	< 2 eV	0

Table 2.1.: Properties of the fermions - the fundamental matter particles in the SM [3].

The fermions interact through the fundamental forces carried by the gauge bosons - the force carriers. The electromagnetic force acts on all charged particles and is mediated by the massless photon γ. It binds the negatively charged electrons and positively charged nucleus (composed of protons and neutrons) together inside an atom. The force responsible for nuclear processes, such as the β decay is the weak force carried by three massive gauge bosons - W^+, W^- and Z. All fermions participate in the weak interaction. Only the weak force acts in flavour[1] changing processes, such as an electron transformed into a muon. The strong force acts on the quarks and at a larger scale binds protons and neutrons inside atomic nuclei to ensure the stability of matter. Eight massless gluons (g) mediate the strong force and are bi-coloured objects composed of a colour and an anticolour combination. Both quarks and gluons with a net colour charge experience the strong interaction. The properties of the gauge bosons in the SM are summarised in Table 2.2. The photon, Z boson and gluons are their own antiparticles while the W^+ is the antiparticle of the W^- boson.

Interaction	Symbol	Mass [GeV]	Charge
Electromagnetic	γ	0	0
Weak (charged)	$W^\pm$	80.387 ± 0.016	± 1
Weak (neutral)	Z	91.188 ± 0.002	0
Strong	g	0	0

Table 2.2.: Properties of the gauge bosons - the force carriers in the SM [3].

[1] Types of different fermions are also known as flavours.

The spin-0 Higgs boson of mass around 125 GeV [3] completes the list of particles in the SM. It is the excitation of the Higgs field responsible for the mass generation of the fundamental particles, as discussed in Subsection 2.1.1.

2.1.1. Electroweak Interaction

The electromagnetic and weak forces are two entirely separate phenomena but can be unified and treated as two aspects of a single force. In 1979, S. L. Glashow, A. Salam and S. Weinberg [12–15] succeeded in unifying the electromagnetic and weak interactions into a single theoretical description, called the electroweak theory or the GSW theory.

As per the individual symmetry groups of the two interactions mentioned earlier, the combined symmetry group of the electroweak interaction is given as $U(1)_Y \otimes SU(2)_L$. The respective conserved quantum numbers associated with the $SU(2)_L$ and $U(1)_Y$ symmetry groups are the three-component weak isospin I and the hypercharge Y. Three components of the weak isospin operator can be expressed in terms of the three Pauli matrices (σ^a, $a = 1, 2, 3$). The hypercharge Y relates I_3 (third component of the weak isospin operator I) and the electrical charge Q through the Gell-Man-Nishijima relation as $Q = I_3 + \frac{Y}{2}$ [20]. Since the Pauli matrices do not commute, the $SU(2)_L$ group is a *non-abelian* group whereas the $U(1)_Y$ group is *abelian*. The index L in the $SU(2)_L$ group indicates the fact that only the left-handed[2] fermions (right-handed antifermions[3]) interact via the weak interaction. Therefore, in this approach, the fermion fields can be split into left and right-handed components. The left-handed fermions form doublets in the weak isospin space ($I = \frac{1}{2}$ and $I_3 = \pm\frac{1}{2}$) while the right-handed fermions are singlets ($I = I_3 = 0$).

Gauge fields represent interactions in the SM. For the electroweak interaction, the gauge fields - B_μ and W^a_μ ($a = 1, 2, 3$) are introduced that belong to the $U(1)_Y$ and $SU(2)_L$ symmetry groups respectively. The Higgs field, a complex scalar doublet Φ with four degrees of freedom, is additionally included [21, 22].

The covariant derivatives describing the electroweak interactions between the gauge fields and the left and right-handed fermion fields $\psi_{\mathrm{R/L}}$ as well as between the scalar Higgs doublet Φ can be written as

$$D_\mu \psi_L = \left(\partial_\mu - i\frac{g}{2} W^a_\mu \sigma^a - i\frac{g'}{2} B_\mu Y\right)\psi_L \tag{2.1}$$

$$D_\mu \psi_R = \left(\partial_\mu - i\frac{g'}{2} B_\mu Y\right)\psi_R \tag{2.2}$$

$$D_\mu \Phi = \left(\partial_\mu - i\frac{g}{2} W^a_\mu \sigma^a - i\frac{g'}{2} B_\mu Y\right)\Phi \tag{2.3}$$

[2] The handedness of a fermion is an intrinsic property, called *chirality* which is equal to *helicity* (projection of a particle's spin on its direction of motion) for massless particles.

[3] For the neutrinos, no right-handed equivalence is observed yet.

In the above set of equations, g and g' refer to the respective coupling constants of the $SU(2)_L$ and $U(1)_Y$ gauge groups. σ^a denote the Pauli matrices. The interactions between the gauge fields with the left and right-handed fermion fields are different as the W^a_μ fields only act on the left-handed fermions or right-handed antifermions. The B_μ gauge field, on the other hand, does not differentiate the fermions depending on their handedness. The Lagrangian describing the electroweak sector including the Higgs field is given as

$$\begin{aligned}
\mathcal{L}_{\text{ewk}} = \; & - \frac{1}{4} W^a_{\mu\nu}.W^{\mu\nu}_a - \frac{1}{4} B_{\mu\nu}.B^{\mu\nu} && (2.4)\\
& + \sum_{\text{fermions}} \left(i\bar{\psi_L} D_\mu \gamma^\mu \psi_L + i\bar{\psi_R} D_\mu \gamma^\mu \psi_R \right) && (2.5)\\
& + (D^\mu \Phi)^\dagger (D_\mu \Phi) - V(\Phi) && (2.6)\\
& - \sum_{\text{fermions}} \lambda_f \left(\bar{\psi_L} \Phi \psi_R + \bar{\psi_R} \bar{\Phi} \psi_R L \right) && (2.7)
\end{aligned}$$

Equation 2.4 in the electroweak Lagrangian includes the kinetic energy terms of the W^a_μ and B_μ fields. Under the gauge invariance, W^a_μ and B_μ gauge field tensors can be expressed as

$$\begin{aligned}
W^a_{\mu\nu} &= \partial_\mu W^a_\nu - \partial_\nu W^a_\mu - g\epsilon^{abc} W^b_\mu W^c_\nu && (2.8)\\
B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\nu && (2.9)
\end{aligned}$$

The bi-linear term $-g\epsilon^{abc} W^b_\mu W^c_\nu$ in the W^a_μ field tensor (Equation 2.8) contains the totally antisymmetric Levi-Civita tensor ϵ^{abc} and is responsible for the self-interaction of the W^a_μ gauge fields. It is due to the *non-abelian* nature of the $SU(2)_L$ group that allows the W^a_μ fields to couple with each other whereas the *abelian* $U(1)_Y$ group prohibits the self-interaction of the B_μ field. The Lagrangian describing the self-interaction of the W^a_μ gauge fields in Equation 2.4 by employing the expression of the W^a_μ field tensor can therefore be written as

$$\begin{aligned}
\mathcal{L}^{\text{int}}_{\text{gauge}} = -\frac{1}{4} W^a_{\mu\nu}.W^{\mu\nu}_a = \; & - \frac{1}{2} g\epsilon^{abc} \left(\partial_\mu W^a_\nu - \partial_\nu W^a_\mu \right) W^{b\mu} W^{c\nu} && (2.10)\\
& - \frac{1}{4} g^2 \epsilon^{abc} \epsilon^{apq} W^a_\mu W^b_\nu W^{p\mu} W^{q\nu} && (2.11)
\end{aligned}$$

The gauge fields W^a_μ and B_μ do not correspond directly to physically observed particle fields. However, linear combinations of these four fields (W^a_μ and B_μ) can be constructed to produce the physical fields observed in nature. The states W^1_μ and W^2_μ mix to form the charged vector bosons ($W^\pm_\mu$) whereas the photon (A_μ) and Z boson (Z_μ) fields can be represented by the mixing of B_μ and W^3_μ states. These physically observed gauge boson fields are

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp iW_\mu^2\right) \tag{2.12}$$

$$Z_\mu = -\sin\theta_W B_\mu + \cos\theta_W W_\mu^3 \tag{2.13}$$

$$A_\mu = \cos\theta_W B_\mu + \sin\theta_W W_\mu^3 \tag{2.14}$$

where $\theta_W = \tan^{-1}(\frac{g'}{g})$ is the weak mixing angle, also referred to as the Weinberg angle, which is a measure for the mixing between the W_μ^a and B_μ terms. The physical fields (Equations 2.12 - 2.14) can now be employed further in the self-interaction Lagrangian term $\mathcal{L}_{\text{gauge}}^{\text{int}}$ (Equations 2.10 and 2.11) to identify the two triple (WWZ, $WW\gamma$) and four quartic ($WWWW$, $WWZZ$, $WWZ\gamma$ and $WW\gamma\gamma$) gauge boson couplings in the SM. These Triple Gauge Boson Couplings (TGCs) are of particular interest in this thesis and are discussed in details in Section 2.5.

The second term (Equation 2.5) of the electroweak Lagrangian describes the interaction between the fermion and gauge fields.

So far, Equations 2.4 and 2.5 of the electroweak Lagrangian as discussed above evidently describe the interactions between the fermions and electroweak gauge bosons as well as the gauge bosons' self-interaction. However, these can only explain massless gauge bosons and fermions that contradicts our observations (Table 2.1 and 2.2). R. Brout, F. Englert and P. Higgs in 1970 [21, 22] proposed a potential solution to this perplexity of theoretically massless gauge bosons and fermions. The proposed mechanism of introducing the Higgs field and generating masses of the fundamental particles, is referred to as the Brout-Englert-Higgs (BEH) mechanism.

The terms in Equation 2.6 of the electroweak Lagrangian represent the interaction of the gauge fields with the Higgs field along with the Higgs field's self-interaction. The Higgs potential $V(\Phi)$ has a non-vanishing vacuum expectation value v which spontaneously breaks the $SU(2)_L \otimes U(1)_Y$ symmetry and gives masses to the gauge bosons in the SM. The masses of the $W^\pm$ and Z bosons are obtained by absorbing three out of the four degrees of freedom of the Higgs field Φ. The fourth degrees of freedom of the Higgs field manifests itself as the massive spin-0 Higgs boson, the excitation of the Higgs field. The masses of the Higgs and the gauge bosons in the SM generated by the BEH mechanism are

$$m_W = \frac{gv}{2}, \quad m_Z = \frac{v}{2}\sqrt{g^2 + g'^2}, \quad m_\gamma = 0, \quad \text{and} \quad m_H = \sqrt{2\lambda}v$$

where λ is the self-coupling of the Higgs boson. The interaction of the gauge fields with the Higgs field further introduces the trilinear (WWH, ZZH) and quadrilinear ($WWHH$, $ZZHH$) couplings as well as cubic (HHH) and quartic ($HHHH$) self-couplings of the Higgs boson.

Finally, the last term of the electroweak Lagrangian (Equation 2.7) presents the Yukawa-type interaction between the fermion (lepton and quark) fields and the Higgs field that gives masses to the fundamenal fermions as

$$m_f = \frac{\lambda_f v}{2}$$

where λ_f refers to the Yukawa coupling for each fermion f.

2.1.2. Strong Interaction

The interaction between the particles carrying a non-vanishing colour charge (quarks and gluons) is the strong interaction. The field theory Quantum Chromodynamics (QCD) describes this interaction and is based on the symmetry group $SU(3)_C$. The index C refers to the colour charge, the conserved quantum number of this symmetry group. Eight gluon fields are introduced to describe the interaction between the quarks and gluons and also the interaction between the gluons themselves. The *non-abelian* nature of this $SU(3)_C$ gauge group allows the gluons to interact with each other. The self-interaction of the gluons leads to two important characteristics of the QCD theory, referred to as the *colour confinement* and the *asymptotic freedom*.

The *colour confinement* indicates the fact that the particles with a net colour charge are not experimentally observed. Figure 2.1a illustrates this mechanism for a meson consisting of a quark-antiquark ($Q\bar{Q}$) pair. If the quark and the antiquark in the $Q\bar{Q}$ pair are about to be separated, a "colour-tube" is formed by the self-interaction of the gluons between them. The energy contained inside this "colour-tube" increases with the distance between Q and $\bar{Q}$. If the colour field inside the "colour-tube" acquires sufficient energy, a light quark-antiquark ($q\bar{q}$) pair is created. The final state now consists of two colour-neutral mesons ($Q\bar{Q} \rightarrow Q\bar{Q} + q\bar{q} \rightarrow Q\bar{q} + q\bar{Q}$). This mechanism is also known as the string breaking where the QCD string between the initial quark Q and the antiquark $\bar{Q}$ is broken due to the light $q\bar{q}$ pair creation.

The strong coupling constant α_s determines the strength of the strong interaction and shows a dependency on the energy scale of the interaction of interest. At high energies, corresponding to smaller distances and higher momentum transfer Q^2 exchanged in an interaction, the α_s decreases asymptotically. In the limit $Q^2 \rightarrow \infty$, the α_s vanishes and the quarks behave as bare quasi-free particles. This effect is referred to as the *asymptotic freedom*. The energy dependence of α_s as shown in Figure 2.1b, makes it a running coupling constant. Therefore, at high energies (short distances), the QCD is applied as a perturbative theory of the strong coupling constant α_s. For energy lower than approximately 200 MeV [23], the coupling constant becomes large enough to create new quarks and gluons due to the *colour confinement* as explained before and the perturbative QCD is no longer be applicable there.

The calculation of physical observables, such as the cross section of a process, described by the Feynman diagrams involving at least one quark or gluon loop in QCD, might lead to divergent results while integrating over the internal loop momenta. It is due to the perturbative theory that computes these observables which is only effective for a certain energy range. The

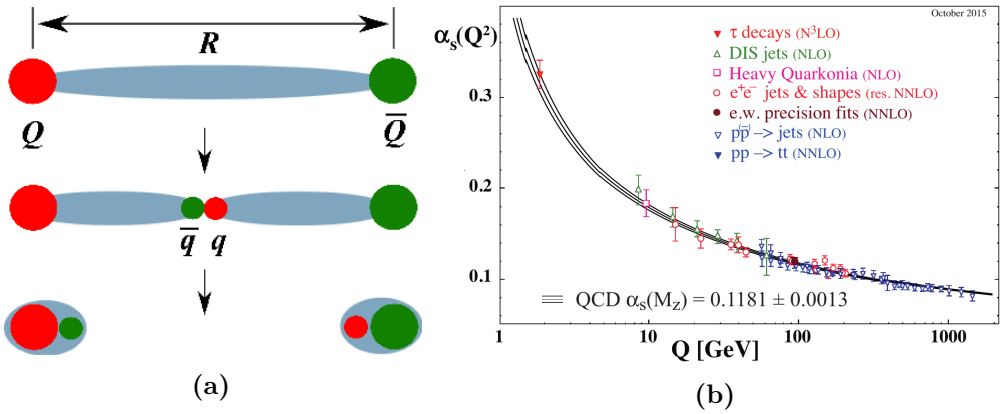


Figure 2.1.: Graphical illustration of string breaking in QCD (Figure 2.1a), taken from Reference [24] and the world average of α_s measured in 2015 as a function of energy scale Q [25] (Figure 2.1b).

contribution from effects outside this energy range can also direct to divergent results. These divergences can be removed by renormalising the coupling constant α_s from its bare value to the experimentally measured physical value at the energy scale of the process of interest. Therefore, a new energy scale, known as the *renormalisation scale* (μ_R) [16, 26], is introduced at which the physical coupling constant $\alpha_s(\mu_R^2)$ is evaluated. The choice of μ_R is arbitrary and is further discussed in Section 2.2.

The strong interaction governs the proton-proton collisions at the hadron colliders where both the *colour confinement* and *asymptotic freedom* play a crucial role as discussed in the following.

2.2. Proton-Proton Collisions

A proton is a composite particle consisting of three valence quarks (uud) that determine its associated quantum numbers. It also contains sea quarks - virtual quarks and antiquarks produced by gluon splitting into quark-antiquark pairs. Partons (quarks and gluons) are held inside the proton by the strong force, mediated by gluons.

In the high energy proton-proton (pp) collisions, hard scattering occurs between the individual partons inside the protons. It is a direct consequence of the *asymptotic freedom* as discussed earlier. The perturbative QCD theory can describe this short distance hard scattering process at a high precision. In addition to the hard scattering process, long distance soft (low energy) interactions also happen between the rest of the partons via the exchange of gluons, which in turn create new quarks and gluons and interact further. Therefore, the evolution of the pp collisions consists of a short distance hard interaction well predicted by the perturbative QCD as well as long distance soft interactions where the perturbative QCD is not applicable. The latter is commonly referred to as the non-perturbative effects. The factorisation theorem [27]

separates these short and long distance effects at an arbitrarily chosen energy scale, referred to as the factorisation scale μ_F.

The perturbative and non-perturbative evolution phases of the pp collisions at the hadron colliders are discussed below.

Partonic Cross Section

The partonic cross section $\hat{\sigma}_X$ of the hard scattering interaction describes probability or cross section for the production of a process in a pp collision. It is in general proportional to the matrix element squared $|\mathcal{M}_{if}|^2$. The matrix element includes all physical information of the transition between the initial state i and the final state f along with the couplings and spins of the involved particles in the hard scattering process. It is calculated from the Lagrangian of the process using the Feynman rules. The cross section of a process $AB \to X$ in a pp collision is therefore calculated using $\hat{\sigma}_X$ as

$$\sigma(AB \to X) = \sum_{a,b} \int dx_a dx_b f_{a/A}(x_a, \mu_F^2) f_{b/B}(x_b, \mu_F^2) \times \hat{\sigma}_X(ab \to X) \tag{2.15}$$

where the two incoming protons A and B collide and the parton a (b) carries the momentum fraction $x_{a(b)}$ of the incoming proton A (B). The Parton Distribution Function (PDF) $f_{a(b)}$ is the probability of the parton a (b) to carry the momentum fraction $x_{a(b)}$. Using the perturbative expansion of $\hat{\sigma}_X$ up to a given order N, Equation 2.15 can be formulated as

$$\sigma(AB \to X) = \sum_{a,b} \int dx_a dx_b f_{a/A}(x_a, \mu_F^2) f_{b/B}(x_b, \mu_F^2) \times \left[\hat{\sigma_0} + \alpha_s(\mu_R^2)\hat{\sigma}_1 + ... + \alpha_s^N(\mu_R^2)\hat{\sigma}_N\right]_{ab\to X} \tag{2.16}$$

where $\hat{\sigma}_i$ is the partonic cross section at a specific order i of the perturbation theory. The perturbative series includes an infinite number of terms which is impossible to calculate in reality. Therefore, the theoretical cross section calculation is pursued to a finite number of terms. The cross section calculated at the first term of the perturbative series, refers to the leading order (LO), up to the second term indicates the next-to-leading order (NLO) cross section calculation and so on. The renormalisation and factorisation scales are denoted as μ_R and μ_F respectively. The more orders in the perturbative expansion of α_s are included in the cross section calculation, the less dependent the calculation becomes on the chosen μ_R and μ_F scales. A common choice of these energy scales is $\mu_R = \mu_F$ that are often fixed to the scale of the hard interaction, such as the invariant mass of the final state particles. The dependencies on these chosen scales introduce an uncertainty, called scale uncertainty in the cross section calculation. If the process of interest involves resonant production, such as W or Z boson, the

subsequent decay of these resonances are also taken into account in the partonic cross section by considering the respective branching ratios[4].

The Parton Distribution Function (PDF) defines the probability of a certain parton carrying the momentum faction x of the initial proton at the momentum transfer scale Q^2. The Q^2 evolution of the PDFs is analytically calculable by the DGLAP equations [28–31]. However, the PDF dependency on the momentum fraction x cannot be calculated analytically and therefore determined from experimental datasets. Several working groups provide the PDFs obtained by fitting a significantly large amount of data from multiple experiments. Figure 2.2 shows the x dependence of the PDF of individual partons at two different values of Q^2. For low Q^2, the proton is dominated by valence quarks (u, d) while at low x, sea quarks and gluons populate the proton.

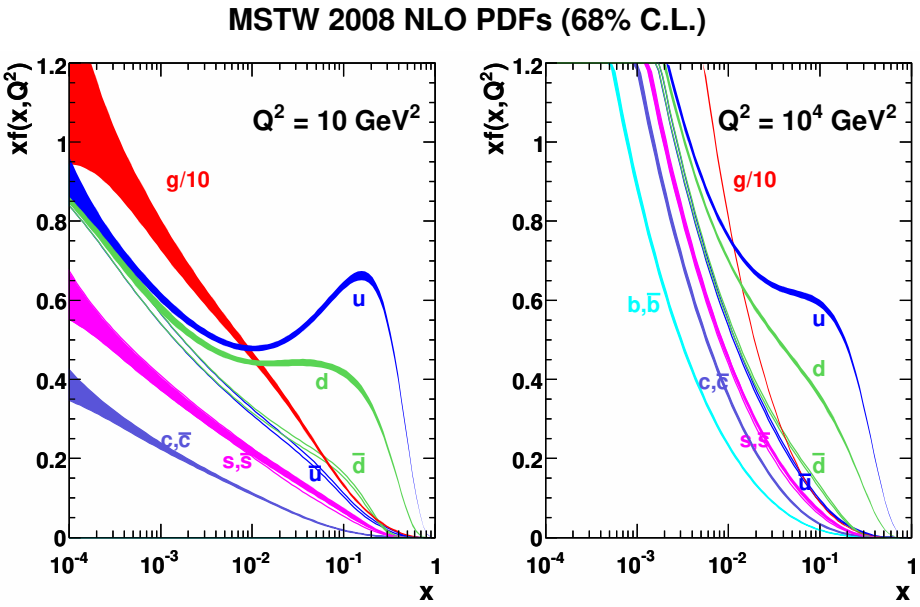


Figure 2.2.: Parton distribution functions as a function of momentum fraction x for two momentum transfer scales $Q^2 = 10$ GeV^2 (left) and $Q^2 = 10^4$ GeV^2 (right), provided by the PDF fitting group MSTW [32].

Parton Shower

The hard scattering in a pp collision is described by the partonic cross section at a fixed order in the perturbative QCD, a given PDF set and certain choices of the factorisation and renormalisation scales. The colour field connects the strongly interacting coloured particles involved in this hard interaction. As the particles move away from each other, the energy of the colour field increases and hence additional gluon radiation, gluon splitting occur resulting in

[4]The branching ratio determines how likely a particle decay in a certain decay mode.

many gluons and quark-antiquark pairs. This parton shower starts at the high energy scale and continues downwards up to the low energy where the perturbation theory breaks down and the *colour confinement* confines the shower of partons into hadrons.

Hadronisation

The produced partons in the showering process cannot exist freely due to the *colour confinement*. Therefore, the partons are recombined to form colour-neutral hadrons. This hadronisation process can be described by phenomenological models and two such common models are the *lund-string model* [33] and *cluster model* [34].

Underlying Events

The underlying events refer to the additional interactions not originating from the hard scattering in a pp collision. Such effect consists of fragmentation of the proton remnants as well as multiple scattering in which more than one parton pairs interact. These additional interactions mostly occur at a low Q^2 scale and are also described by phenomenological models.

The Monte Carlo (MC) simulation of physics processes in the pp collisions considers all these effects as mentioned above and is described in the following section.

2.3. Monte Carlo Generation

Monte Carlo (MC) simulation is necessary for all physics processes that contribute to an analysis. The simulated samples predict the expected sensitivity of the detector exposed to various physics processes and estimate the relative contribution of each process in the real recorded data. To facilitate this, MC production undergoes several computing stages where both the hard scattering perturbative and soft non-perturbative interactions in a pp collision are simulated. A physics event is first generated by MC event generators that calculate the matrix element for a given physics process. The matrix element involves the initial hard scattering process of a fixed number of incoming or outgoing particles using the perturbative QCD calculation at leading order (LO) or next-to-leading order (NLO) in power of α_s. The respective PDFs of the contributing partons in the hard scattering process describe their kinematics. The showering of the outgoing particles and their hadronisations are modelled in the next step. Finally, the underlying events are simulated and overlaid. The generated event is then propagated through the virtual ATLAS detector infrastructure using the Geant4 (GEometry ANd Tracking) [35, 36] software that simulates the particle-detector interactions or hits. In the sensitive material of the detector, these hits are collected to construct the full simulated event eventually.

At different parts of this analysis, MC events are considered at various evolution phases of the pp collisions. The parton-level refers only to the hard scattering interaction before going through the parton shower and hadronisation processes. The particle-level or truth-level indicates the phase where the hard scattering and non-perturbative effects are already simulated in the event. The detector-level or reconstruction-level points to the events reconstructed in the detector.

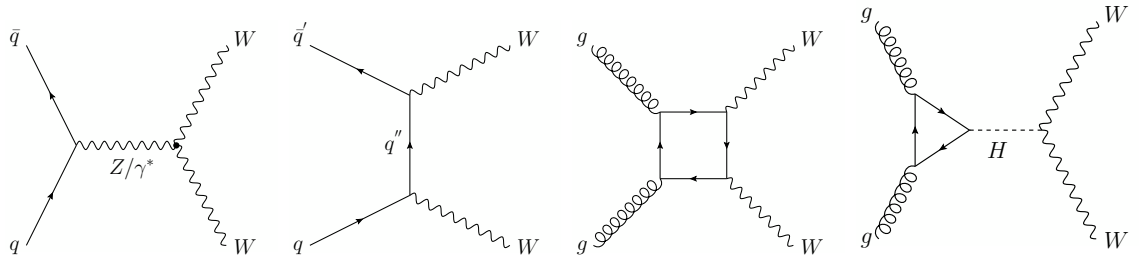
2.4. W^+W^- Production

The W^+W^- (simply denoted as WW) production is the most abundant diboson production process in the SM. At the LHC, the $pp \to WW$ final state is predominantly produced by the quark-antiquark ($q\bar{q}$) interaction.

The $q\bar{q} \to WW$ production mechanism involves the s-channel annihilation and the t-channel scattering processes and the respective leading order (LO) Feynman diagrams are depicted in Figures 2.3a and 2.3b. The s-channel diagram demonstrates the production of the WW final state through the exchange of a Z/γ^* boson and involves the triple gauge boson vertices $WW\gamma$ and WWZ. In the t-channel production mechanism, the two W bosons in the final state are the result of the scattering between the initial quark and antiquark. The WW production by quark-gluon (qg and $\bar{q}g$) interaction starts contributing at next-to-leading order $\mathcal{O}(\alpha_s)$ in the perturbative QCD and has a minimal contribution in the total WW production cross section in comparison with the $q\bar{q} \to WW$ process. The $q\bar{q}$, qg and $\bar{q}g \to WW$ processes are combinedly denoted as $q\bar{q} \to WW$ process in the following.

Besides the $q\bar{q}$ interaction, the W boson pair can also be produced in the gluon fusion that contributes only via loop diagrams appearing at next-to-next-to-leading order $\mathcal{O}(\alpha_s^2)$ in the perturbative QCD. The gluon fusion mediated via a quark loop can produce the WW final state as presented in the Feynman diagram 2.3c. This process is referred to as the non-resonant $gg \to WW$ process. However, gluon fusion can also produce an intermediate Higgs boson which eventually creates a W boson pair in its subsequent decay. This $gg \to H \to WW$ process as shown in the Feynman diagram 2.3d is known as the resonant Higgs boson production.

In the following sections, the WW process is referred to as the combination of the three subprocesses - $q\bar{q} \to WW$, non-resonant $gg \to WW$ and resonant $gg \to H \to WW$ unless explicitly stated. The relative contributions of the $q\bar{q} \to WW$, non-resonant and resonant gg productions to the total WW cross section are 87%, 5% and 8% respectively. The cross section predictions of these three subprocesses are discussed in details in Chapter 6.



(a) s-channel. **(b)** t-channel. **(c)** Non-resonant. **(d)** Higgs boson production.

Figure 2.3.: Leading order Feynman diagrams of the WW production via $q\bar{q}$ and gg interactions. The s-channel diagram (Figure 2.3a) produces the W boson pair through the exchange of a Z/γ^* boson and involves the triple gauge boson vertices - $WW\gamma$ and WWZ. The t-channel production (Figure 2.3b) shows the two W bosons scattered off the initial quark-antiquark pair. Figure 2.3c and 2.3d are the respective Feynman diagrams of the W boson pair production in gluon fusion through a quark loop and an intermediate Higgs boson.

The *non-abelian* structure of the electroweak theory permits the electroweak gauge bosons to self-interact as discussed in Section 2.1.1. The self-interaction of the gauge bosons directs to two important features of the WW production process. The energy dependence of the cross section is discussed below whereas Section 2.5 describes the sensitivity of the gauge bosons' self-couplings to physics beyond the SM.

The partonic cross sections of the s and t-channel $q\bar{q} \to WW$ production processes vary with the centre-of-mass collision energy $\hat{s}$ as $\sim \frac{1}{\hat{s}} \ln \hat{s}$ [37]. The individual cross section calculation for these two channels grows to infinitely large value with $\hat{s}$. But as soon as both channels are combined, a negative interference acts between the channels. This interference effect introduces a cancellation that constraints the combined cross section to a finite value as per expectation. This cancellation itself is the outcome of the *non-abelian* structure of the electroweak gauge theory and only persists if all the existing couplings including the self-couplings of the gauge bosons in the SM are as predicted.

The interference between the s and t-channels can be visualised in Figure 2.4 where the measured $e^+e^- \to WW$ cross section in the electron-positron (e^+e^-) collisions with the LEP experiments [38] is compared with the theoretical predictions. The WW production mechanisms in the e^+e^- collisions are analogous to the $q\bar{q} \to WW$ process as shown in the Feynman diagrams 2.3a and 2.3b after replacing the $q\bar{q}$ pair with the e^+e^- pair. The t-channel scattering of the $e^+e^- \to WW$ process occurs through the exchange of a neutrino ν_e whereas the s-channel annihilation process produces the WW final state via the exchange of a γ or Z boson and involves the $WW\gamma$ and WWZ vertices. The theoretical predictions are shown only for the t-channel, in the absence of WWZ vertex and combining the s and t-channel contributions. The cross section dampening after adding both the contributing channels is the first clear evidence of the *non-abelian* nature of the electroweak theory.

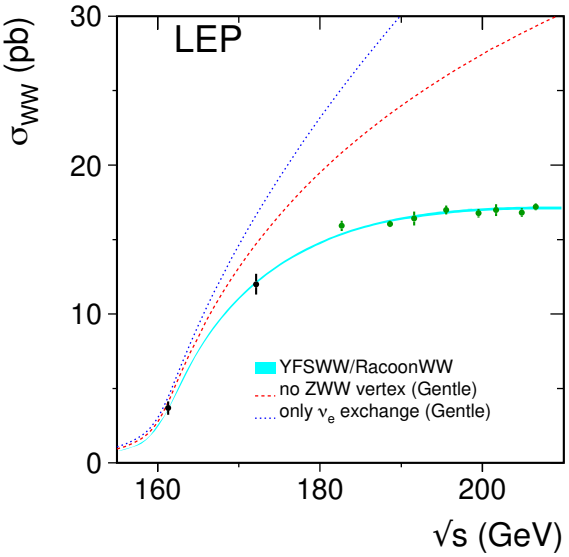


Figure 2.4.: Measured $e^+e^- \to WW$ cross section in the e^+e^- collisions at the LEP experiments. The dashed lines represent the theoretical cross section predictions only for the t-channel i.e. in the absence of $WW\gamma$ and WWZ vertices (blue), in the absence of WWZ vertex (red) and the combined s and t-channel contributions (cyan) [38].

2.5. Anomalous Triple Gauge Boson Couplings

The WW production process via $q\bar{q}$ interaction in the s-channel (Figure 2.3a) involves two Triple Gauge Boson Couplings (TGCs) - $WW\gamma$ and WWZ that have fixed values in the SM. However, these two couplings can deviate from their predicted SM values as experimental results [38] do not highly constraint them. Any deviations in those couplings would cause an enhanced production cross section of the WW process at higher energies. The deviations of the TGCs from their SM expectations are referred to as the anomalous Triple Gauge Boson Couplings (aTGCs) that can be probed in this WW production process.

An effective Lagrangian [39] with a most general form under the condition that the charge and parity are individually conserved, can be written to describe these TGCs as

$$\frac{\mathcal{L}_{eff}^{WWX}}{g_{WWX}} = -i\left[g_1^X X^\mu \left(W^-_{\mu\nu}W^{\mu+} - W^+_{\mu\nu}W^{\nu-}\right) + k_X W^+_\mu W^-_\nu X^{\mu\nu} + \frac{\lambda_X}{M_W^2} X^{\mu\nu} W^{\rho+}_\nu W^-_{\rho\mu}\right] \quad (2.17)$$

where g_1^X, K^X and λ^X with X being either a γ or a Z boson, are the newly introduced coupling parameters. The overall coupling constants g_{WWX} are $g_{WW\gamma} = e$ and $g_{WWZ} = e\cot\theta_W$ with e being the electrical charge of the positron. The θ_W refers to the weak mixing angle.

The self-interaction Lagrangian describing the triple and quartic gauge boson couplings in the SM are constructed in Equations 2.10 and 2.11 in the form of the W^a_μ fields. However, the physical gauge boson fields as obtained in Equations 2.12 - 2.14 can be inserted into Equation 2.10 to express the TGCs in terms of the physically observed gauge bosons W, Z and γ. This can be written as

$$\begin{aligned}\mathcal{L}^{\text{int}}_{\text{gauge}}|_{\text{TGCs}} = &- ig\left[W^-_{\mu\nu}W^{\mu+} - W^+_{\mu\nu}W^{\nu-}\right]\left(\cos\theta_W Z^\mu + \sin\theta_W A^\mu\right) \\ &- \frac{ig}{2}\left(\cos\theta_W Z_{\mu\nu} + \sin\theta_W A_{\mu\nu}\right)\left[W^{\mu-}W^{\nu+} - W^{\mu+}W^{\nu-}\right] \end{aligned} \tag{2.18}$$

The comparison of the effective Lagrangian (Equation 2.17) and the SM Lagrangian for TGCs (Equation 2.18) leads to the SM values of the new coupling parameters as $g_1^X = K^X = 1$ and $\lambda^X = 0$. The deviations of these parameters representing the anomalous couplings (aTGCs) are commonly parametrised as

$$\Delta g_1^X = g_1^X - 1 \tag{2.19}$$

$$\Delta k^X = k^X - 1 \tag{2.20}$$

$$\lambda^X \tag{2.21}$$

where all six independent aTGC parameters vanish in the SM. The electromagnetic gauge invariance furthermore fixes Δg_1^γ to zero and leaves only five independent aTGC parameters - $\Delta g_1^Z, \Delta k^Z, \lambda^Z, \Delta k^\gamma$ and λ^γ.

Moreover, two simplifying assumptions motivated by the $SU(2)_L \otimes U(1)_Y$ gauge invariance restrict the number of independent aTGC parameters from five to three. These assumptions, referred to as the "LEP scenario", are as follows

$$\begin{aligned}\Delta k^Z &= \Delta g_1^Z - \Delta k^\gamma \tan^2\theta_W \\ \lambda &= \lambda^\gamma = \lambda^Z\end{aligned}$$

In the SM these anomalous coupling parameters are exactly zero to avoid tree-level unitarity violation. So, if there exists no beyond the SM (BSM) physics, these anomalous couplings parameters would violate the unitarity bounds. This violation can be prevented by introducing $\sqrt{\hat{s}}$ (invariant mass of the WW system) dependent Form Factor (FF) that would modify the anomalous coupling k as

$$k(\sqrt{\hat{s}}) = \frac{k(0)}{(1 + \frac{\sqrt{\hat{s}}}{\Lambda^2_{\text{FF}}})^2} \tag{2.22}$$

where Λ_{FF} is the energy scale where new physics affecting the anomalous couplings would occur. The upper bounds on the aTGCs can be obtained at this Λ_{FF} [40]. The "no form factor scenario" ($\Lambda_{\mathrm{FF}} \to \infty$) is also often used in the limit setting.

An alternative parametrisation of the aTGCs parameters is described by the Effective Field Theory (EFT) approach that obeys the $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ gauge invariance and does not introduce an arbitrary form factor. This approach is demonstrated in Reference [41].

The aTGC parameters relating the WWZ coupling can also be studied in the WZ production process while the $W\gamma$ production can probe the anomalous couplings at the $WW\gamma$ vertex.

In this analysis, five independent aTGC parameters ($\Delta g_1^Z, \Delta k^Z, \lambda^Z, \Delta k^\gamma$ and λ^γ) are investigated without any imposed constraints or form factor (implying $\Lambda_{\mathrm{FF}} = \infty$).

Chapter 3.

The LHC and the ATLAS Detector

This chapter introduces the experimental setup to study the WW process at the ATLAS detector, one of the four multipurpose particle detectors situated at the Large Hadron Collider (LHC). Section 3.1 gives a general overview of the LHC accelerator and describes the condition of the proton-proton (pp) beams collided at the ATLAS detector in 2015. A description of the detector and its subsystems built to detect a large number of particles produced in the pp collisions is provided in Section 3.2.

3.1. The Large Hadron Collider

The Large Hadron Collider (LHC) [42], world's most powerful hadron-hadron collider is located at the European Organization for Nuclear Research (CERN) facility in Geneva, Switzerland. It accelerates two counter-rotating proton as well as the heavy ion (Pb) beams inside a ring of 26.7 km circular circumference. The LHC contains four major interaction points at which the two beams are brought to collision.

Multiple pre-accelerators accelerate the protons up to an energy of 450 GeV before they are injected into the LHC ring. The protons are inserted into the ring as bunches where each bunch contains about 10^{11} protons. A bunch spacing of 25 or 50 ns separates two consecutive bunches. So far, proton beams are accelerated inside the ring to the maximum beam energy of 6.5 TeV corresponding to the maximum centre-of-mass collision energy of $\sqrt{s} = 13$ TeV.

The LHC delivers an extremely high instantaneous luminosity $\mathcal{L}$ that is proportional to the rate of proton-proton interactions and therefore is directly correlated to the data volume recorded at the detectors. It is quoted in units of $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and evaluated as

$$\mathcal{L} = f \frac{n_1 n_2}{4\pi\sigma_x\sigma_y} \tag{3.1}$$

where f is the frequency with which the number of protons n_1 and n_2 in beam 1 and 2 collide. σ_x and σ_y denote the horizontal and vertical spread of the particle beams. The total integrated

luminosity L over a data-taking period is therefore given as

$$L = \int \mathcal{L} dt \tag{3.2}$$

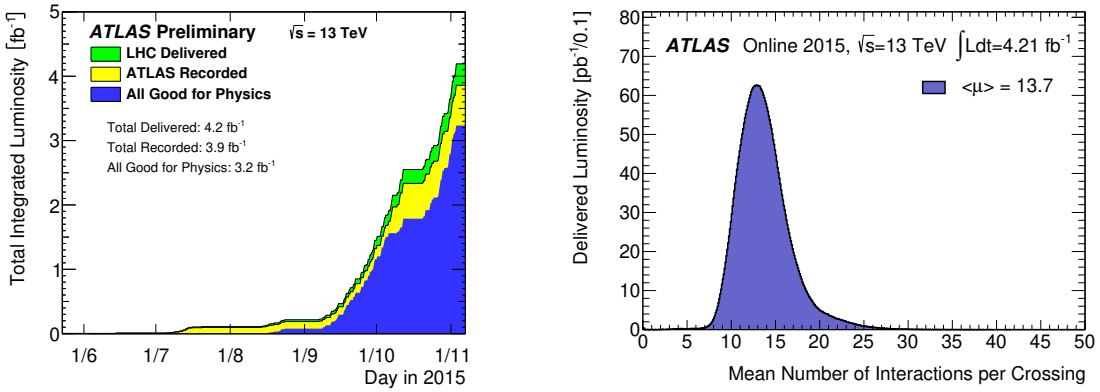
It is related to the total number of interactions N produced by a given physics process with cross section σ as

$$N = \sigma L \tag{3.3}$$

The high instantaneous luminosity causes the number of proton-proton interactions in a single collision to be more than one which is referred to as pileup. However, due to small bunch spacing, additional proton-proton interactions in a collision may also occur from neighbouring (previous or next) bunches. These pileup interactions mostly consist of low energy soft non-perturbative physics processes that accompany one energetic hard-scattered interaction of interest.

3.1.1. Beam Conditions in 2015

After a two-year-long shutdown (LS1 - Long Shutdown I), the LHC has been operated at a centre-of-mass energy of $\sqrt{s} = 13$ TeV in 2015 and its peak instantaneous luminosity reached a value of 5.0×10^{33} cm^{-2}s^{-1} [43]. Figure 3.1a shows the total integrated luminosity delivered by the LHC and recorded by the ATLAS detector with 25 ns bunch spacing in 2015. A total luminosity of 3.2 fb^{-1} is used for physics analyses due to the requirement that all detector subcomponents should be fully functional. The average number of interactions per collision is presented in Figure 3.1b.



(a) **(b)**

Figure 3.1.: Breakdown of the total integrated luminosity delivered by the LHC and collected by the ATLAS detector in 2015 (Figure 3.1a) and the average number of interactions per collision (Figure 3.1b) [43].

3.2. The ATLAS Detector

The ATLAS detector [44] is designed to study proton-proton collisions at the LHC and is constructed to be forward-backwards and cylindrically symmetric with respect to the interaction point. It comprises several subsystems wrapped in layers around the central beam pipe. The innermost subsystem closest to the beam pipe is the Inner Detector (ID) - a tracking system immersed in a 2 T solenoidal magnetic field that measures the momentum of charged particles by the magnetic deflection of their trajectories. The electromagnetic and hadronic calorimeters surround the ID to provide a precise measurement of particle energies. The outermost subsystem is the Muon Spectrometer (MS) - another tracking system to determine muon trajectories. It is embedded inside three superconducting toroidal magnets. Each detector subsystem is segmented into one barrel and two identical end-cap units. The barrel unit surrounds the beam pipe as a concentric cylindrical layer while the end-caps are mounted vertically. A schematic overview of the ATLAS detector is shown in Figure 3.2 and the major subsystems are highlighted.

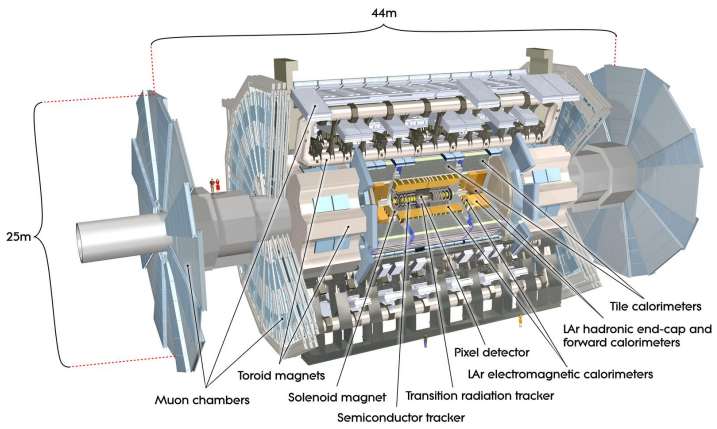


Figure 3.2.: Schematic view of the ATLAS detector [44].

Coordinate System

The ATLAS experiment uses a right-handed coordinate system with its origin at the nominal interaction point at the centre of the detector. The positive x-axis points towards the centre of the LHC ring, the positive y-axis directs vertically upwards and the positive z-axis is along the beam direction. The cylindrical geometry of the detector makes it convenient to work with polar coordinates. The azimuthal angle ϕ is measured from the positive x-axis in the transverse x-y plane whereas polar angle θ is measured from the positive z-axis in y-z plane.

In the high energy particle collisions, the rapidity y is preferred over θ as the difference in rapidity Δy remains Lorentz-Invariant under longitudinal boosts along the beam direction. It is defined as

$$y = \frac{1}{2}\ln\left[\frac{E + p_z}{E - p_z}\right] \tag{3.4}$$

where E is the particle's energy and p_z corresponds to the z-component of its three-momentum vector $\vec{p}$. However, it is difficult to measure the energy and momentum of highly relativistic particles. In such a scenario, another useful quantity pseudorapidity $\eta = -\ln\left(\tan\frac{\theta}{2}\right)$ is considered that only depends on the polar angle θ. In the high relativistic limit with the approximation that the mass of a particle is negligible, the pseudorapidity becomes equivalent to the rapidity. The distance in the η-ϕ plane is defined as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

Along the longitudinal beam direction, the initial momentum of the colliding partons inside the proton is not known and hence the conserved physics quantities[5], such as transverse momentum p_{T}, transverse energy E_{T} and transverse mass m_{T} are measured in the transverse x-y plane.

3.2.1. Inner Detector

The Inner Detector (ID) [45] is placed closest to the interaction point to register charged particles produced in a collision and precisely measures their momenta up to $|\eta| < 2.5$. A solenoidal magnetic field of 2 T parallel to the beam axis acts throughout the ID that bends charged particles depending on their electrical charge. The ID comprises three independent and complementary subsystems - the pixel detector, the Semiconductor Tracker (SCT) and the Transition Radiation Tracker (TRT). Each subsystem further contains several layers of detecting materials placed one after another. Figure 3.3 shows the R-z cross-sectional view of the ID subsystems.

Pixel Detector

The pixel detector is the innermost ID subsystem surrounding the beam pipe that provides the highest resolution for determining the trajectories (tracks) of charged particles. It is instrumented with four layers of silicon pixel sensors. Each pixel layer contributes to a three-dimensional (R, ϕ, z) space point measurement (hit) as a charged particle passes through it. The innermost pixel barrel layer - the Insertable B-Layer (IBL) [47] has been installed as an extra tracking layer during LS1 (Figure 3.3). The insertion of the IBL improves the overall tracking efficiency of the pixel detector even in the presence of a large number of particle tracks produced in higher luminosity and centre-of-mass energy collisions. The IBL is complemented by three outer pixel

[5] The conserved quantities of a particle of mass m, energy E and three-momentum vector $|\vec{p}| = (p_x, p_y, p_z)$ produced in a pp collision, can be expressed in terms of polar angle θ as
$p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2} = |\vec{p}|\sin\theta$, $p_z = |\vec{p}|\cos\theta$, $E_{\mathrm{T}} = E\sin\theta$ and $m_{\mathrm{T}} = \sqrt{m^2 + p_{\mathrm{T}}^2}$.

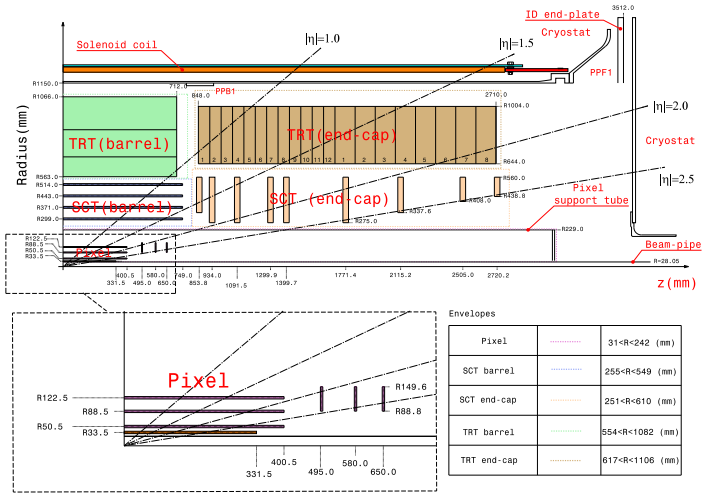


Figure 3.3.: *R-z* cross-sectional view for one half of the ID from the interaction point ($z = 0$). The top panel lists all ID subsystem layers. The bottom panel shows a magnified view of the pixel detector including the Insertable B-Layer (IBL), the innermost pixel barrel layer inserted during LS1 [46].

layers arranged as three cylindrical barrel layers and three perpendicular disks in each end-cap region.

The position measurement in IBL sensors of size 50×250 µm^2 has an intrinsic azimuthal (R-ϕ plane) resolution of 8 µm and axial (z) resolution of 40 µm in the barrel region only [48]. The outer three pixel layers comprise sensors of size 50×400 µm^2 and provide 10 µm azimuthal resolution in R-ϕ (z-ϕ) plane and 115 µm axial resolution in z (R) direction over the barrel (end-caps) unit [48].

Semiconductor Tracker

The Semiconductor Tracker (SCT) is segmented into four barrel layers and nine end-cap disks of silicon microstrip sensor modules. Each module consists of four strip sensors, two at the top and two at the bottom. The top and bottom pair of sensors on alternate sides of a single module are placed at a relative angle of 40 mrad to record full three-dimensional position measurement. Strip sensors are designed to have 17 µm azimuthal and 580 µm axial resolutions in the barrel and end-cap units [48]. The SCT complements the pixel tracking system by providing at least four additional hits for better momentum resolution of the tracks.

Transition Radiation Tracker

The Transition Radiation Tracker (TRT) is constructed with drift tubes (straws), each with a diameter of 4 mm and filled with an active xenon based gas mixture. The areas, affected by the gas leakage before LS1, were flushed with argon gas as a substitute for the expensive xenon

during this LS1 shutdown phase. The inter-gap between the tubes are filled with polypropylene polymer that acts as transition radiation material. Charged particles crossing the straw tubes ionise the TRT gas and produce two-dimensional hits in the azimuthal plane with an intrinsic resolution of 130 μm per straw [48]. The TRT with its coarser resolution than the pixel and SCT detectors is expected to add $\sim$36 hits to the precision tracking within $|\eta| < 2.0$. In addition to the ionisation hits, low energy transition radiation photons are also produced as the particles cross the polymer material. The probability of emitting these photons depends on the relativistic Lorentz factor $\gamma = \frac{E}{m}$ of the incoming particle and hence the emission is more likely to happen for electrons than the charged pions. This characteristic provides a means to separate electrons from charged pions.

3.2.2. Calorimeter Systems

The ATLAS detector uses sampling calorimeters [49] with alternating layers of an active sensing material and a passive absorber. It is designed to absorb the entire energy of electromagnetic (EM) and hadronic showers produced by the interaction of incident particles with the detector material. It has two primary systems. The electromagnetic calorimeter (commonly referred to as the EM calorimeter) directly surrounding the ID provides the energy measurement of EM objects, such as electrons and photons. The hadronic calorimeter surrounding the EM calorimeter is optimised for hadronic particles stemming from the hadronisation of produced quarks and and are usually collimated in jets.

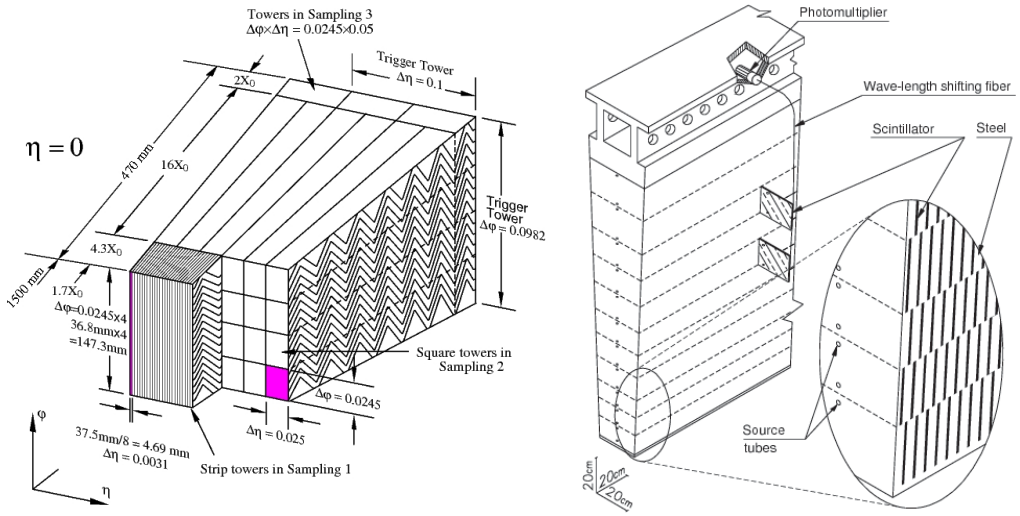
Electromagnetic Calorimeter

The EM calorimeter utilises liquid argon (LAr) as the sensing material, lead plates as an absorber and accordion-shaped kapton electrodes as the readout system. This accordion geometry gives the possibility of full coverage in ϕ with no azimuthal cracks. The barrel part of the EM calorimeter is divided into two halves at $z = 0$ and covers up to $|\eta| < 1.475$. Two end-caps extend to a pseudorapidity range of $1.375 < |\eta| < 3.2$.

The energy is measured in cells with varying η, ϕ granularity in the three longitudinal layers of the EM calorimeter as shown in Figure 3.4a for a barrel segment. The front strip layer with very fine η granularity can separate showers initiated by electrons or photons from neutral pions by resolving the impact points of the two closely collimated photons from the pion decay. The middle layer includes even finer granularity in ϕ and is crucial for energy measurement of EM objects. The third layer with the coarsest granularity receives the tail of the EM showers. Energy depositions in these three layers contribute to the final energy measurement. A single layer LAr presampler detector ($|\eta| < 1.8$) is placed in front of the EM calorimeter to estimate the energy loss of electrons and photons before entering into the strip layer. The EM calorimeter measures the energy of EM objects with a designed energy resolution of $\frac{\sigma_E}{E} = \frac{10\,\%}{\sqrt{E}} \oplus 0.7\,\%$ [44].

Hadronic Calorimeter

The hadronic calorimeter is designed to fully stop the particle showers initiated by hadronic objects in the detector and measures their energy. It comprises the barrel tile calorimeter ($|\eta| < 1.7$) and the LAr hadronic end-caps ($1.5 < |\eta| < 3.2$) with increasingly coarser granularity. The tile calorimeter makes use of steel-plate absorbers and plastic scintillators as sensing material while the LAr hadronic end-caps use copper and liquid argon as the absorber and sensing element respectively. Figure 3.4b shows a schematic view of the tile calorimeter where the scintillators are coupled with wavelength-shifting fibres that transmit the scintillator signals to the photomultiplier tubes for readout. The barrel and end-cap hadronic calorimeters are designed to provide an energy resolution of $\frac{\sigma_E}{E} = \frac{50\,\%}{\sqrt{E}} \oplus 3\,\%$ [44]. The LAr forward calorimeters measure the energy of both electromagnetic and hadronic objects in the forward regions ($3.1 < |\eta| < 4.9$).



(a) **(b)**

Figure 3.4.: Longitudinal three-layer segmentation of the EM calorimeter with their respective η, ϕ granularities (Figure 3.4a) and schematic view of a barrel segment of the tile calorimeter (Figure 3.4b) [44].

3.2.3. Muon Spectrometer

The Muon Spectrometer (MS) [50] is fully integrated within three large air-core toroids to measure muon trajectories bent by magnetic deflection. One barrel ($|\eta| < 1.0$) and two end-cap ($1.4 < |\eta| < 2.7$) toroids include eight coils each aligned radially around the beam pipe. The region between $1.0 < |\eta| < 1.4$ is the transition region of the barrel and end-cap detector components. The air-core minimises the material density that the muons must traverse. This reduces the possibility of multiple scattering and hence improves the tracking resolution. In the

transition region, magnetic deflection occurs due to overlapping of the barrel and end-cap fields and the performance has been significantly improved by inserting additional detectors during LS1.

The high precision muon tracking system involves two different detector technologies to determine momentum and position of muon tracks in the bending (R-z) plane. Precise muon tracking mainly comes from the Monitored Drift Tube (MDT) chambers comprising three to eight layers of drift tubes filled with pressurised argon gas. Muon hits are detected in the ionised gas with a single-hit spatial resolution of 80 µm. Due to the high radiation levels in the region of $2.0 < |\eta| < 2.7$, the MDT chambers are replaced with the Cathode Strip Chambers (CSCs), multiwire proportional chambers with higher rate capability and improved spatial resolution of 60 µm.

The MS is capable of triggering muon tracks within the region $|\eta| < 2.4$ using precision trigger chambers - the Resistive Plate Chambers (RPCs) and the Thin Gap Chambers (TGCs). Both RPCs ($|\eta| < 1.05$) and TGCs ($1.05 < |\eta| < 2.4$) complement the tracking chambers by measuring η and ϕ coordinates of the muon tracks.

3.2.4. Trigger System

With a nominal bunch spacing of 25 ns, the LHC is designed to deliver a collision rate of 40 MHz. However, due to limitations in technologies and computing resources, only a fraction of the total event rate can be recorded. The ATLAS Trigger and Data Acquisition System (TDAQ) [51] records events produced in the pp collisions at a manageable readout rate of 1 kHz. Besides, it preserves events interesting for physics analyses. The trigger system involves a two-level trigger selections [52] - a hardware-based Level-1 (L1) trigger and a single software-based High-Level Trigger (HLT).

The L1 trigger uses customised electronics and is instrumented directly in the detector. To identify objects with high momentum, it processes raw energy deposits (energy cluster) in the calorimeter systems and muon information coming from the muon trigger chambers. The L1 trigger defines Region-of-Interest (RoI) around each identified object passing a specific p_{T} threshold using the raw detector information and sends the decision (to select or reject an event) to the HLT. It reduces the bunch crossing rate of 40 MHz to the readout event rate of 100 kHz with a decision latency of less than 2.5 µs.

The HLT trigger further reduces the readout event rate to 1 kHz. It performs a full event reconstruction by processing every subsystem information within the identified RoI. Only events passing the HLT are permanently stored and used in physics analyses.

Reconstructing objects during trigger-level event processing is referred to as the *online* reconstruction as opposed to the *offline* reconstruction that is carried out on the recorded events as discussed in Chapter 4.

Chapter 4.

Event Reconstruction

A recorded event is processed through a sequence of *offline* algorithms that reconstructs the full event and classifies the physics objects inside it.

The classification of physics objects inside an event begins with constructing trajectories or tracks of the charged particles by combining the hits in the ID. The reconstructed tracks are then associated with the interaction points, called vertices from which the tracks originate. The track and vertex reconstructions are discussed in Section 4.1.

Tracks matched with their respective vertices are finally used to reconstruct the physics objects inside an event with or without being combined with the energy deposits (clusters) in the calorimeter systems or the MS hits. A series of optimisation is performed on the reconstructed objects to improve their quality. Sections 4.2 - 4.4 describe these *offline* object reconstruction algorithms and criteria to select physics objects, such as electrons, muons and jets that are interesting for this analysis. Section 4.5 discusses the reconstruction of missing transverse momentum to infer indirectly measured neutrinos that pass through the detector without leaving any trail.

If more than one object is built from a single track or energy cluster, an ambiguity resolution is performed that prioritises a specific object and discards the others. This procedure is known as the overlap removal and detailed in Section 4.6.

Not all data events reconstructed by the algorithms as mentioned earlier, participate in this analysis. Some of them might be affected by high detector noise or reconstruction inefficiencies. Therefore, a data quality cross-check is performed to remove these poor events from data to ensure the selection of good quality events. This is achieved using the Good Runs List (GRL) selection as presented in Section 4.7.

Moreover, the same set of algorithms (Sections 4.1 - 4.6) as applied to the data are also used to reconstruct and select physics objects in the simulated MC events. The efficiencies of each of the object selection steps (Section 4.1 - 4.4) are compared between the data and MC samples. The differences in data and MC efficiencies are applied for the corrections of MC samples to replicate the data responses. Section 4.8 describes the applied MC corrections and

the associated object selection uncertainties. These corrections and uncertainties influence the MC predictions for physics processes that eventually contribute to this analysis.

4.1. Tracks and Vertices

Track reconstruction uses two sequential algorithms [53] that combine the hits produced by a charged particle in the ID subsystems into a helical trajectory.

The *inside-out* algorithm begins with creating a preliminary seed track from a set of three hits starting from the inner pixel layers towards the SCT. A hit entering the seed must be compatible with the expected hit position in the detector based on the momentum and impact parameters[6] estimated from the track curvature in the uniform magnetic field. A combinatorial *Kalman filter* approach [54] further builds the track candidate from the chosen seed by adding consecutive hits in the remaining pixel and SCT layers. In this pattern recognition process, multiple track candidates can be formed if there exists more than one hit in a single layer compatible with the seed track. In such a scenario, a track satisfying quality criteria, such as higher hit multiplicities is given priority. A global χ^2 track fit reconstructs the track segment in the silicon (the pixel and SCT) detectors. The silicon track segment is later extended to the TRT by adding successive TRT hits. Eventually, a final track fit obtains the full reconstructed track combining the information of all ID layers.

Reconstructing a track with the *inside-out* algorithm relies entirely on the seed track formation. Secondary charged particles coming from kaon decays or photon conversions might have few or no hits in the pixel and SCT layers to form the seed track. Hence, an alternative back-tracing *outside-in* algorithm is considered. It starts with reconstructing the track segment in the TRT and is later extended inwards by adding consecutive hits in the SCT and pixel detectors. The TRT track segments with no associated SCT or pixel hits are recorded as the TRT-only tracks.

To limit the number of reconstructed tracks in the full ID coverage ($|\eta| < 2.5$), only tracks with $p_{\mathrm{T}} > 0.4$ GeV and minimum seven hits in the silicon detectors are further processed [55].

Tracks are extrapolated back to the beamline to reconstruct the vertices in an event. A *vertex finding algorithm* [56] is used to create a vertex seed by looking at the global maximum in the z distribution of the extrapolated tracks. An iterative χ^2 fit is performed on the vertex seed and the tracks around it, that refits the seed position. If a track is displaced by less than seven standard deviations (7σ) from the seed position, it is assigned to that vertex candidate. Otherwise, it is removed from the input list of the fit and considered to seed a new vertex. The process continues until no additional vertices or unassociated tracks are found in the event. In this procedure, an event can have more than one vertex and each vertex requires at least

[6] The longitudinal and transverse distance of the extrapolated track to the beamline from the interaction point.

two associated tracks. The vertex that contains the maximum $\sum p_{\mathrm{T}}^2$ of all associated tracks is chosen as the primary vertex.

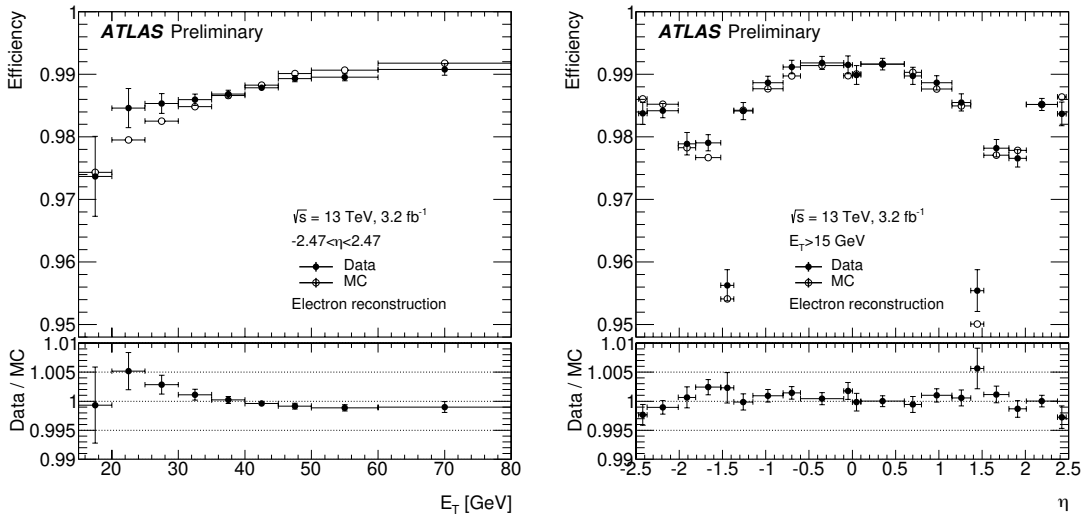
4.2. Electrons

Electron reconstruction [57] combines the EM calorimeter energy cluster with compatible ID track within $|\eta| < 2.47$. An incoming EM particle deposits energy into many cells in the three longitudinal EM calorimeter layers. All these cells affected by the incoming particle is combinedly referred to as the energy cluster.

To perform the *offline* cluster reconstruction [58], the EM calorimeter is segmented into a grid of $N_\eta \times N_\phi = 200 \times 256$ towers. The size of the towers is equivalent to the EM calorimeter middle layer granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. The deposited cell energies are measured inside these towers and for a single particle, the total energy cluster can be distributed to several neighbouring towers. The algorithm searches for cluster seeds with a minimum transverse energy E_{T} of 2.5 GeV using a sliding window of size 3×5 towers. Simultaneously, reconstructed tracks with $p_{\mathrm{T}} > 0.5$ GeV are extrapolated to the points of impact in the EM calorimeter middle layer. A track is then matched with a cluster seed to reconstruct an electron candidate, provided the distance between the track's point of impact and the seed cluster barycentre is within $|\Delta\eta| < 0.05$. If multiple tracks get combined to a single seed, a selection is made based on the maximum number of pixel and SCT hits or minimum distance from the seed position. Electron tracks are then refitted using an improved tracking algorithm *Gaussian-Sum filter* (GSF) [59] to account for the energy losses due to the bremsstrahlung process. The electron cluster energy is finally recomputed using a 3×7 (5×5) towers in the barrel (end-caps) region of the EM calorimeter. The total electron energy is evaluated by adding the cluster energy with the estimated energy deposited around the cluster (lateral leakage), outside the cluster in the hadronic calorimeter (hadronic leakage) and in the ID and presampler detector.

The reconstruction efficiency is defined as the fraction of all reconstructed electron clusters that are associated with ID tracks to form electron candidates and further satisfy specific track quality criteria. It is measured in $Z \to ee$ events selected in data and MC sample as a function of both transverse energy E_{T} and pseudorapidity η as shown in Figure 4.1. In the central region of the detector ($|\eta| < 2.47$), the reconstruction efficiency is $\sim$98% except in the transition region of $1.37 < |\eta| < 1.52$ (Figure 4.1b). The efficiency in this transition region, also known as the crack region, drops significantly due to a large amount of material placed at the boundaries between the barrel and end-caps of the EM calorimeter.

The reconstructed electrons include the contribution of genuine signal electrons from heavy resonance decays, such as $W \to e\nu, Z \to ee$ along with background contaminations. Background electrons originate from hadronic jets or semi-leptonic decays of heavy quarks (b quarks) that exhibit similar detector responses as the signal electrons. The electrons from photon conversions



(a) **(b)**

Figure 4.1.: Data/MC comparisons for electron reconstruction efficiency as the function of electron E_{T} (Figure 4.1a) and η (Figure 4.1b). The shown uncertainty includes both statistical and systematic uncertainties associated with the efficiency measurements [60].

are also treated as background as they are not originated from the signal processes of interest. A multivariate likelihood-based technique [57] is used to distinguish the signal electrons from the associated backgrounds. It utilises the shape and size of the EM showers produced by electrons in the EM calorimeter along with the track property information. A likelihood discriminant is constructed to simultaneously analyse each of these properties and identify an electron being signal or background. Imposing selections on this likelihood discriminant introduces three electron identification operating points - *Tight*, *Medium* and *Loose*. Each operating point corresponds to a certain signal selection efficiency and background rejection. The *Loose* criterion has the highest signal efficiency but lowest background rejection as it is mostly dominated by jet backgrounds containing u, d, s and c quarks. The *Tight* operating point, on the other hand, provides the lowest signal efficiency but the highest background rejection. This analysis primarily selects *Tight* electrons whereas in certain areas, it does require *Medium* electrons as well. Figure 4.2 shows the combined electron reconstruction and identification efficiencies measured in $J/\psi \to ee$ and $Z \to ee$ events for low E_{T} ($7 < E_{\mathrm{T}} < 15$ GeV) and high E_{T} ($E_{\mathrm{T}} > 15$ GeV) electrons respectively.

Identified signal electrons are mostly isolated while the background electrons tend to align with the presence of nearby objects. An isolation requirement provides further background suppression by applying selections on nearby energy deposits or tracks surrounding the electron candidate. This can be achieved by both calorimeter and track-based measurements. The calorimeter-based isolation is measured by subtracting the electron cluster energy from the total energy deposited within a cone size of $\Delta R = 0.2$ around the electron cluster. The track-based isolation follows the same procedure. It is calculated by subtracting the p_{T} of the primary electron track (p_{T}^{e})

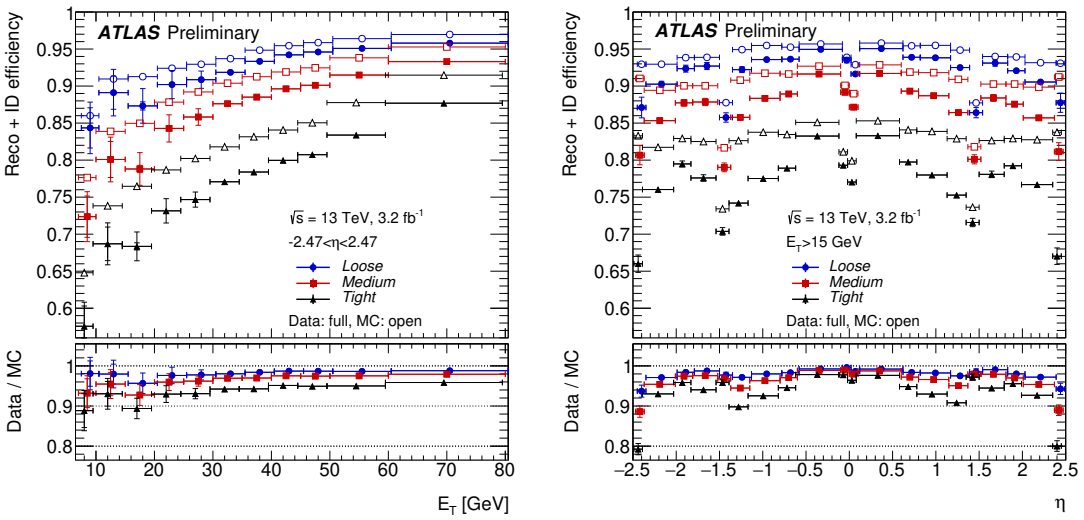


Figure 4.2.: Combined electron reconstruction and identification efficiencies in data and MC as a function of electron E_{T} (left) and η (right). The shown uncertainty includes both statistical and systematic uncertainties [57].

from the total p_{T} of all tracks in a variable cone size of $\Delta R = \min(0.2, 10\,\mathrm{GeV}/p_{\mathrm{T}}^{e})$. These two measurements are finally combined into isolation operating points optimised for specific efficiencies to meet the need for different analyses. In this analysis, the *Gradient* isolation operating point is chosen. The efficiency of reconstructed and identified electrons from $Z \to ee$ decay process to satisfy the *Gradient* isolation selection is $\sim 90\,(99)\%$ for electrons with p_{T} $> 25\,(60)$ GeV [57].

4.3. Muons

The muon candidates are formed by combining the reconstructed track segments in the ID and MS tracking systems. The procedure to reconstruct tracks as detailed in Section 4.1 is applied to construct the ID track segment. The MS track segment uses a combination of hits in multiple layers of the muon chambers. The muon reconstruction algorithm [61] classifies the muons into the following four classes where each class uses different subdetector (ID, calorimeters or MS) responses to determine the muon trajectory.

- **Combined (CB):** A track segment in the MS is extrapolated and combined with an ID track segment to obtain the full muon trajectory.
- **Segment-tagged (ST):** An ID track segment is recognised as originating from a muon if associated with at least one local track segment in either MDT or CSC when extrapolated. This class is optimal for low p_{T} muons as they might not travel through all the MS detecting layers due to the low momentum.

- **Calorimeter-tagged (CT):** An ID track segment is matched with calorimeter energy deposits to reconstruct a muon object. The deposited energy in such case must be compatible with the expected energy deposits from minimum ionising particles, such as muons. This category is beneficial to muons within $|\eta| < 0.1$ where the MS offers almost no coverage.
- **Extrapolated (ME):** In the region of $2.5 < |\eta| < 2.7$, the ID provides no coverage. Therefore, the MS track segment cannot be matched with any ID track segment but extrapolated close to the interaction point.

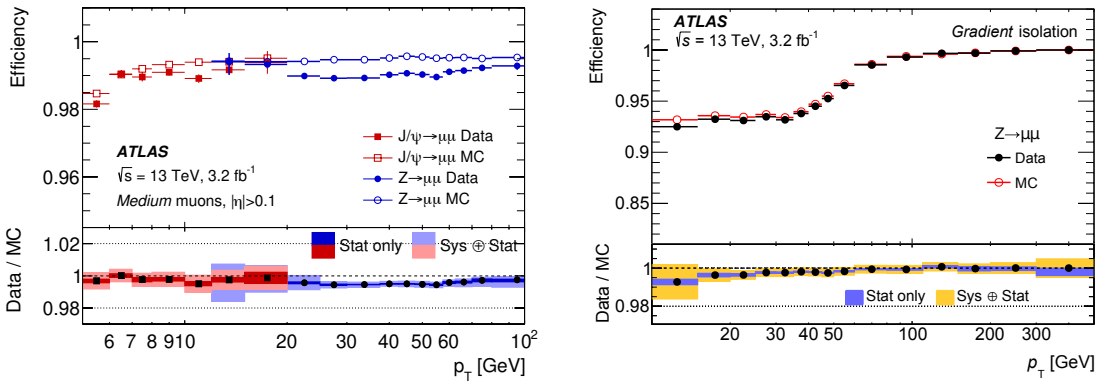
An identification criterion is imposed on the reconstructed muon candidates to select the signal muons produced by the W, Z or the Higgs boson decays and reject the associated background objects. Such background objects primarily originate from the decay of light-hadrons (pions, kaons). Three identification operating points - *Loose*, *Medium* and *Tight* are constructed to offer a definite signal efficiency and background rejection by applying a different level of selections on track and calorimeter variables. This analysis uses muons selected with *Medium* operating point that accepts only the CB and ME reconstruction classes.

Figure 4.3a shows the data and MC reconstruction efficiencies of *Medium* muons with $|\eta| > 0.1$ as a function of muon p_{T}. The efficiencies are obtained from $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ events selected in both data and MC for low and high p_{T} muons respectively.

The signal muons are mainly well separated or isolated from other particles. But, the background muons from the semi-leptonic hadronic decays are embedded inside the hadronic jets. The isolation criterion measures the activity of nearby objects around the muon candidate and provides additional background rejection. The muon isolation operating points combine both calorimeter and track-based isolation measurements performed using a similar procedure as applied to the electrons. The calorimeter-based measurement is defined by adding the total deposited transverse energy in a cone size of $\Delta R = 0.2$ around the muon candidate and subtracting the energy deposited by the muon itself. The track-based measurement calculates the p_{T} of all tracks and removes the muon track p_{T} contribution (p_{T}^{μ}) from it in a cone size of $\Delta R = \min(0.3, 10\,\mathrm{GeV}/p_{\mathrm{T}}^{\mu})$ around the muon track. A *Gradient* isolation operating point is chosen to be applied on the muon candidates used in this analysis. The data and MC efficiencies measured in $Z \to \mu\mu$ events for *Medium* muons to satisfy the *Gradient* isolation requirement is presented in Figure 4.3b.

4.4. Jets

Jets are originated from the hadronisation of initial quarks, gluons and subsequent hadronic decays of W/Z bosons ($W \to q\bar{q}$, $Z \to q\bar{q}$). They produce collimated particle showers and deposit energy in the calorimeter systems.



(a) **(b)**

Figure 4.3.: Data/MC comparisons of both the reconstruction (Figure 4.3a) and isolation (Figure 4.3b) efficiencies of *Medium* muons as a function of muon p_{T}. The uncertainties on the efficiencies presented in the top panels are only statistical while the bottom panels show the ratio of data to MC efficiencies including both statistical and systematic uncertainties [61].

The jet reconstruction process searches for topologically connected clusters in the calorimeters. These clusters combine neighbouring cells with significant energy deposits above the background noise level. An iterative *anti-k_t* [62] algorithm uses the topological clusters as inputs to build the jet candidates. This algorithm iterates over every pair of topological clusters i, j and computes the distance $d_{i,j}$ between them. It also determines the distance d_i of each cluster i from the beam axis. Both these distance parameters - $d_{i,j}$ and d_i are defined as

$$d_{i,j} = min(\frac{1}{k_{t,i}^2}, \frac{1}{k_{t,j}^2})\frac{\Delta R_{i,j}^2}{R^2}, \qquad d_i = \frac{1}{k_{t,i}^2} \tag{4.1}$$

where $k_{t,i}$ is the transverse momentum of the i^{th} cluster and R is the reference radius parameter that roughly determines the size of the jet candidate. If $d = \min(d_{i,j}, d_i)$, the minimum of the two distance parameters mentioned above, becomes equal to $d_{i,j}$, then the clusters i and j are merged and the process continues. In case of $d = d_i$, the i^{th} cluster is identified as a complete jet cluster and removed from the algorithm process. The process continues until all clusters are identified as jet clusters. Jet clusters are then combined with the associated ID tracks to complete the reconstruction process. In this analysis, reconstructed jets with reference radius parameter $R = 0.4$ are selected.

The reconstructed jet energy is thereafter calibrated using an EM+JES calibration scheme [63] to the true energy of the jet originating from a stable hadronic particle. The energy calibration accounts for detector responses and the effect of multiple pp interactions (pileup) in the same bunch crossing. The average energy of the pileup interactions is subtracted from the measured jet energy for pileup correction. The $\eta-p_{\mathrm{T}}$ dependent corrections are also applied to the jet energy derived from dedicated MC simulation. Further corrections consider the longitudinal and

transverse structure of the jets, inactive areas of the detector as well as jets that are not fully contained in the calorimeters. For jets in data, an additional *in-situ* [64] correction is applied that calibrates the jet energy to account for the difference between the data and MC simulation in jet energy measurement. These calibrations are usually referred to as the Jet Energy Scale (JES) response. The calibrated central jets (jets in the central region of the detector within $|\eta| < 2.4$) with $p_{\mathrm{T}} > 25$ GeV and forward jets (within $2.4 < |\eta| < 4.5$) with $p_{\mathrm{T}} > 30$ GeV are selected in this analysis.

In addition to the uniform pileup correction applied by the EM+JES scheme, pileup jets can originate from QCD effects from a single pileup vertex or multiple vertices. The Jet Vertex Tagger (JVT) technique [65] is used to separate the hard-scattered jets from these pileup jets by extrapolating them back to their primary vertices. The JVT is two-dimensional likelihood discriminant that uses both jet track and vertex information to ensure that most of the selected jets must be associated with their respective primary vertices. In this analysis, a JVT > 0.64 requirement for central jets with $p_{\mathrm{T}} < 50$ GeV is used that provides 92% efficiency in selecting only the hard-scattered jets of interest.

Identification of b-jets

The identification of jets originating from the bottom (b) quarks is essential for many analyses where physics processes involving the top quarks play an influential role. The top quark in such processes predominantly decays into a b quark that eventually hadronises into a b-jet in the detector.

The b-hadrons (containing b quarks) inside the b-jet have longer lifetimes than other hadrons inside the light-jets (coming from u, d, s quarks or gluons) and c-jets (from c quarks). Hence, the b-jet secondary vertex is displaced from the primary vertex of the event. A multivariate *b-tagging MV2c20* algorithm [66] uses accurate track reconstruction, impact parameter, primary and secondary vertex information of reconstructed and calibrated central jets with $p_{\mathrm{T}} > 20$ GeV. It is trained to discriminate b-jets from light-jets or c-jets using MC simulation. The signal sample comprises events with b-jets and the background sample is composed of 20% c-jet and 80% light-jet fractions. Four operating points are constructed by imposing selections on the *MV2c20* output distribution that provide specific b-jet tagging (b-tagging) efficiencies of 60%, 70%, 77% and 85%. Each operating point corresponds to a different rate of c-jet and light-jet rejections [67].

In this analysis, the 85% b-tagging efficiency operating point is chosen to identify the b-jets in an event. Figure 4.4 presents the data and MC efficiencies of this operating point as a function of the jet p_{T}.

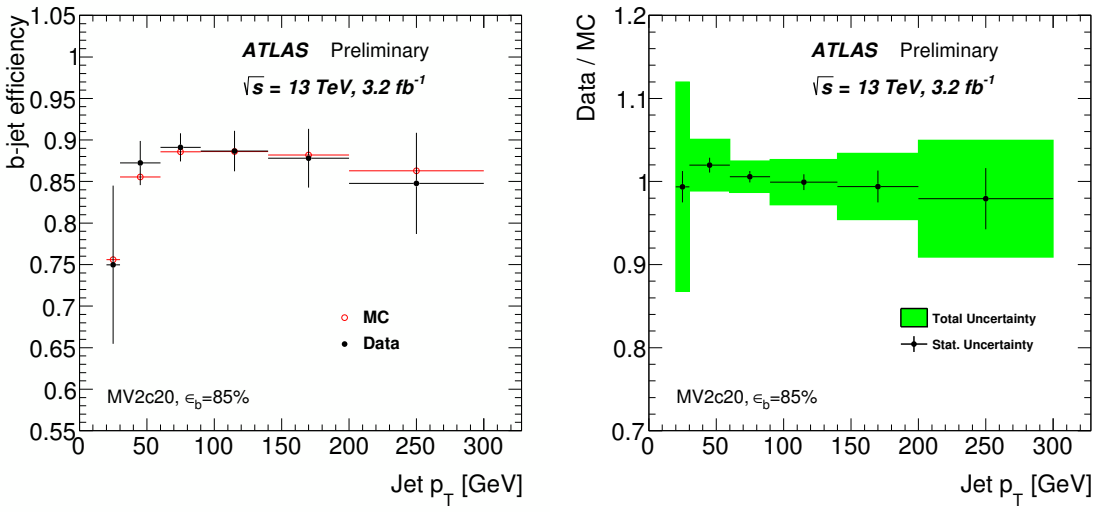


Figure 4.4.: Data/MC efficiency comparisons for 85% b-tagging efficiency operating point (left) as a function of jet p_{T}. The statistical and total uncertainties on this efficiency measurement are shown in the right figure [68].

4.5. Missing Transverse Momentum

In pp collisions, the colliding partons carry no initial momentum transverse to the beam direction. Therefore, the vectorial sum of the transverse momenta of all reconstructed objects in an event should be zero. Any imbalance in the sum is measured as the missing transverse momentum with magnitude $E_{\mathrm{T}}^{\mathrm{miss}}$ which indicates the presence of undetectable neutrinos. It could also suggest the existence of particles as predicted by theories beyond the SM. Objects that are poorly reconstructed or not reconstructed at all as well as lying outside the detector acceptance, also contribute to this $E_{\mathrm{T}}^{\mathrm{miss}}$ term.

Equation 4.2 demonstrates the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation as the negative vectorial sum of the transverse momentum of all reconstructed hard-scattered objects in an event. In this analysis, only electrons, muons and jets are specifically calibrated for the $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement. Other objects, such as photons and hadronically decaying τ leptons in the event, enter the $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement as reconstructed jets. The soft term as introduced in the following equation corresponds to tracks from the primary vertex or calorimeter energy clusters that are not matched with any of the reconstructed objects.

$$E_{x(y)}^{\mathrm{miss}} = -p_{x(y)}^{e} - p_{x(y)}^{\mu} - p_{x(y)}^{\mathrm{jets}} - p_{x(y)}^{\mathrm{soft}} \tag{4.2}$$

The $E_{\mathrm{T}}^{\mathrm{miss}}$ can be measured using calorimeter or track-based measurements. The calorimeter-based term (referred to as $E_{\mathrm{T}}^{\mathrm{miss}}$ from this point onwards) uses cluster energy information. The track-based measurement considers the track p_{T} information of all reconstructed objects and is denoted as $p_{\mathrm{T}}^{\mathrm{miss}}$. The soft term similarly can either be measured using the energy clusters not associated with any reconstructed tracks, known as Calorimeter Soft Term (CST) or be

determined from tracks associated with the primary vertex that are not suited to any clusters, denoted as Track Soft Term (TST). The involved soft term in both $E_{\mathrm{T}}^{\mathrm{miss}}$ and $p_{\mathrm{T}}^{\mathrm{miss}}$ calculations used to measure the neutrino contribution in this analysis, utilises the TST estimation within the ID acceptance and CST estimation in the forward regions outside the ID acceptance.

The $p_{\mathrm{T}}^{\mathrm{miss}}$ is mostly insensitive to high pileup effects due to higher accuracy in the track-vertex association [69]. However, neutral particles with no associated tracks are not included in the $p_{\mathrm{T}}^{\mathrm{miss}}$ calculation. On the other hand, the $E_{\mathrm{T}}^{\mathrm{miss}}$ offers a better response than $p_{\mathrm{T}}^{\mathrm{miss}}$ as $p_{\mathrm{T}}^{\mathrm{miss}}$ by construction does not include neutral particles (unless they convert or decay beforehand) [69].

The $E_{\mathrm{T}}^{\mathrm{miss}}$ in an event is highly sensitive to the momenta mismeasurement of the hard-scattered objects (leptons or jets) if it is aligned to such nearby objects. The probability of selecting such events can be reduced by taking the transverse projection of $E_{\mathrm{T}}^{\mathrm{miss}}$ against the nearby lepton or jet. This projection is referred to as the relative missing transverse momentum ($E_{\mathrm{T,rel}}^{\mathrm{miss}}$) and calculated as

$$E_{\mathrm{T,rel}}^{\mathrm{miss}} = \begin{cases} E_{\mathrm{T}}^{\mathrm{miss}} \times \sin\left(\Delta\phi_{\ell,j}\right) & \text{if } \Delta\phi_{\ell,j} < \pi/2 \\ E_{\mathrm{T}}^{\mathrm{miss}} & \text{if } \Delta\phi_{\ell,j} \geq \pi/2, \end{cases} \tag{4.3}$$

where $\Delta\phi_{\ell,j}$ is the angular difference between the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the nearby lepton or jet.

4.6. Overlap Removal

Electron, muon and jet candidates are reconstructed independently using a combination of tracks and calorimeter clusters. A single energy cluster or track can be used to reconstruct several objects. This ambiguity is resolved using an overlap removal procedure that selects a specific object and discards the others if they share the same track or energy cluster.

Electron-Muon Overlap Removal - If a track is associated with both an electron and a muon within $\Delta R(e,\mu) < 0.02$, the muon candidate is prefered if it satisfies all muon selection criteria. The electron is hence discarded as it cannot have a track segment in the MS.

Electron-Jet Overlap Removal - If a jet and an electron candidate lie within $\Delta R(e,j) < 0.2$, the electron is selected over the jet candidate since the jet does not distinguish between the produced EM and hadronic showers. On the contrary, the electron is removed and the jet persists for $0.2 < \Delta R(e,j) < 0.4$ as the electron reconstructed near the edge of the jet most likely originates from the semi-leptonic decays of the jet constituents.

Muon-Jet Overlap Removal - If a muon is within $\Delta R(\mu,j) < 0.2$ of a jet having less than three associated tracks, the jet gets discarded since the jet is most likely reconstructed from the muon's energy loss in the calorimeters. In case of $0.2 < \Delta R(\mu,j) < 0.4$, the muon candidate is removed and the jet is selected as the muon might come from semi-leptonic decays of particles inside the jet.

4.7. Good Runs List

Recorded data must undergo a data quality cross-check for each *lumiblock* that assures the selected data to be of high quality and interesting for physics analyses. The *lumiblock* is a short interval in a data-taking run (typically 60 seconds) where the data is considered to be stable after verifying that all detector components were fully functional. This selection is achieved by employing the Good Runs List (GRL) that registers all these stable runs. Any recorded event not belonging to the list is removed and considered non-interesting. Moreover, the used GRL determines the total luminosity taken into account in an analysis. The dataset passing the GRL selection used in this analysis corresponds to a total integrated luminosity of $L = 3.16$ fb^{-1}.

4.8. Experimental Uncertainties

Physics processes contributing to this analysis are mostly modelled with simulated MC samples. MC predictions of these processes determine their relative contributions in the collision data. However, to attain a proper data to MC comparison, MC samples must exhibit responses similar to data in the object reconstructions and selections in an event. Therefore, a series of corrections and object selection uncertainties are assigned to the MC samples for proper data compatibility as discussed below.

4.8.1. Monte Carlo Corrections

MC samples are assigned with a given cross section either calculated theoretically or measured experimentally. Predicting the number of events for a specific process requires the used MC sample to be scaled to the total integrated data luminosity according to Equation 3.3.

Several reweighting techniques applied to MC samples, are utilised to replicate the kinematic distributions observed in the collision data.

A pileup reweighting is performed that reweights the MC pileup distribution to reproduce the data pileup profile. In this procedure, an additional MC rescaling factor of 1.16 is applied to the MC to match the number of primary vertices observed in data. This factor is obtained by comparing the primary vertex distributions in both data and MC. This rescaling factor improves the data to MC agreement by recovering the measured (in data) number of inelastic collisions in MC samples.

Furthermore, MC samples are weighted by the data to MC efficiency differences observed in object reconstruction and selection procedures. For electrons and muons, efficiency differences come from the reconstruction, identification and isolation selections as described in Sections 4.2 and 4.3. Scale Factors (SFs) are calculated as the data to MC efficiency ratio for each of the lepton selection steps. The calculated SFs correct the MC efficiencies to replicate the data

responses. For muons, an additional correction related to the efficiency of the track-vertex association is used. Each analysis selects the final state objects by imposing an analysis specific trigger selection. For this analysis, the trigger criteria are applied to the electron and muon candidates. Therefore, trigger SFs, calculated by taking the ratio of data and MC trigger efficiencies of electrons and muons, also correct MC events. The reconstructed jets (Section 4.4) in the MC events are corrected using η-p_{T} dependent correction factors to account for the differences in JES responses between the data and MC. Moreover, the additional MC weights from jet selections are associated with the JVT and b-tagging efficiency measurements.

4.8.2. Detector Systematic Uncertainties

The estimated number of events using MC sample are required to be converted into detector independent measurements. Therefore, any source of uncertainties either related to pp beams affecting the recorded dataset or detector-level object selection must be taken into account. These uncertainties include - the luminosity measurement of the used dataset, pileup reweighting of the simulated MC samples to the observed data and the selected objects from the detector responses. The list of detector systematic uncertainties impacting the MC-based predictions is discussed below.

Luminosity and Pileup

The total integrated luminosity $L = 3.16$ fb^{-1} of the dataset used in this analysis has an assigned uncertainty of 2.1% from the luminosity measurement [70]. This uncertainty is applied independently to all MC samples that are apriori scaled by the total luminosity of the used dataset.

The MC pileup uncertainty originates from the calculation of the additional rescaling factor of 1.16. The MC prediction is varied with $\sim$6% up/down (1.09/1.23) variations of this factor and the deviation from the nominal prediction is assigned as the pileup uncertainty.

Electrons and Muons

The electron and muon systematic uncertainties either originate from the energy calibrations or the SF measurements.

The electron energy measured in data is highly affected by the energy loss of the electrons outside the sensitive calorimeter volume. Therefore, the electron energy scale and resolution are calibrated in MC samples and corrected to match the data. The calibration is performed by comparing the Z peak in $Z \to ee$ events selected in both data and MC [71]. The associated uncertainties to this calibration come from the samples statistics, the choice of the MC generator used in this study, the energy loss in the presampler detector and also interactions of electrons

with the detector materials. The electron energy scale and resolution are measured with approximately 0.5% and less than 1% precision respectively that vary within a range of up to a few per mil depending on the η and p_{T} of the electron candidate. For muons, the momentum scale and resolution are determined likewise but comparing the Z peak in the $Z \to \mu\mu$ events. The muon momentum resolution is measured with a 2.3% precision while the momentum scale has an uncertainty of 0.05% in the central region of the detector [61].

For both electron and muons, measured SFs have their assigned total (statistical and systematic) uncertainties originating from the measurement procedures [57, 61]. For electrons, the SF uncertainties arise from the reconstruction, identification, isolation and analysis specific trigger selections. In the case of muons, for a particular identification operating point - *Medium* used in this analysis, the other selection (reconstruction and isolation) efficiencies are calculated. Hence, the muon uncertainties originate from the reconstruction, isolation and trigger SF measurements. For leptons with $p_{\mathrm{T}} > 25$ GeV, the individual lepton SF has an assigned uncertainty of approximately below 1%.

Jets

The jet selection uncertainties come from the JVT efficiency and calibration of JES and Jet Energy Resolution (JER). An uncertainty of 1% associated with the JVT efficiency measurement is propagated to the JVT SF that corrects the MC samples to the data. The dominant jet uncertainty originates from energy calibrations [63, 72]. The JES uncertainty is governed by the dependencies on the jet flavour compositions, pileup effects and additionally the η-p_{T} dependent correction factors applied to the MC to match the observed data. It is split into a number of nuisance parameters to account for individual systematic effect in the energy scale calibration. For central jets, the JES uncertainty ranges from 6% for jets with $p_{\mathrm{T}} > 20$ GeV to less than 2% for jets with $p_{\mathrm{T}} > 100$ GeV. The energy resolution is measured with an uncertainty of 4% (1%) for jets with $p_{\mathrm{T}} > 20$ (200) GeV.

Jet Flavour Tagging

The jet flavour tagging refers to the identification of jets either originating from b quarks, c quarks or other light-flavoured (u, d, s) quarks or gluons. The b-tagging efficiency and the corresponding misidentification rate are discussed in Section 4.4 where the efficiency to identify a b-jet is measured in both data and MC. The calculated SF (Figure 4.4) is applied to the MC that takes the data to MC efficiency difference into consideration. The misidentification rate is indicated as the probability to tag a jet as a b-jet when it is a c-jet or a light-jet. The same principle is applied to calculate the c-jet and light-jet SFs. The uncertainties on the jet flavour tagging SFs are treated as systematics variations affecting the MC predictions.

$\boldsymbol{E}_{\mathbf{T}}^{\mathbf{miss}}$

The uncertainties on the $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement mainly originate from the associated soft term calculation. An overall uncertainty of $\sim 2\%$ is assigned to the energy scale and resolution calculations of this soft term [73].

The MC prediction is varied by the total up/down uncertainties associated with each object selection or calibration step and the deviation from the nominal prediction is considered as an individual source of detector uncertainties on the prediction.

The reconstructed data and MC events are finally propagated to the analysis measurement as discussed in Chapter 5.

Chapter 5.

Measurement Overview

This chapter discusses the selection of $pp \to WW$ signal candidate events in a defined analysis phase space. The phase space is chosen to achieve a signal enriched region, denoted as the Signal Region (SR). It is optimised to reduce contaminations from processes other than $pp \to WW$, in the following sections referred to as backgrounds. The signature of the signal and contributing background processes that enter the SR are described in Section 5.1.

The MC samples used to model these processes are addressed in Section 5.2. The SR selection utilises kinematic distributions of the final state objects originating from the decays of both W bosons. Section 5.3 discusses the SR event selection used in this analysis. It also presents the contributions of signal and background events predicted by the MC samples in the selected data.

The total selected events after the SR event selection contain a high fraction of WW signal candidates. The remaining background contributions from MC predictions are further determined using dedicated methodologies. Section 5.4 details the estimations of individual background processes and their associated uncertainties in the SR.

5.1. Signal and Background Processes

The signal and background processes entering this analysis selection are presented in the following.

Signal Processes

The $pp \to WW$ production by $q\bar{q}$ and gg interactions as detailed in Section 2.4 are the signal processes contributing to this analysis. The W boson can either decay hadronically, producing jets in the final state or leptonically into a lepton (electron, muon or tau) and its associated neutrino. In this analysis, only the leptonic decay mode of two bosons in the WW final state is considered where one of the bosons decays into an electron and an electron neutrino and the

other W boson produces a muon with an associated muon neutrino. The W boson's decay into tau lepton is not counted as the experimental signal signature of this analysis since the tau lepton decays immediately into other particles either hadronically or leptonically. The electron and muon final state particles must be of opposite electrical charge. Both possible decay modes $e^+\mu^-$ and $e^-\mu^+$ are combined and denoted as the $e\mu$ pair. The $e\mu$ pair can be further split into $e\mu$ and μe channels where the former lepton is recognised as the leading lepton carrying the highest p_{T}. The same flavour di-lepton decay channels ($e^\pm e^\mp$ or $\mu^\pm \mu^\mp$) are not included in this analysis to reduce large background contribution from the Drell-Yan process as discussed in the following section.

Background Processes

Backgrounds contributing to the WW candidate sample are physics processes that exhibit the same $e\mu$ final state signature as the signal. This section describes each of these processes in details.

Top

One of the dominant backgrounds in this analysis comes from processes involving top (t) quarks. The top background includes both top quark pair ($t\bar{t}$) production as well as single top quark production in association with a W boson, referred to as the Wt process. The top quark predominantly decays into a W boson and a b quark where the b quark hadronises into a b-jet in the detector. Both $t\bar{t}$ and Wt processes produce the WW final state in association with mostly b-jets or other light-jets produced in the subsequent decays. The two W bosons further decay into an $e\mu$ pair and enter the analysis selection as background. Discarding events containing additional jets (mainly b-jets) reduces the top background significantly.

Drell-Yan

The Drell-Yan process, referred to as the Z/γ^* production, is another significant background to enter the WW signal selection. The Z/γ^* decays into a same flavour, opposite charge lepton pair ($Z/\gamma^* \to ee/\mu\mu/\tau\tau$). Since the signal selection requires a lepton pair of different flavour in the final state, the contribution from $Z/\gamma^* \to ee/\mu\mu$ process is highly suppressed. The Drell-Yan background mainly comes from the $Z/\gamma^* \to \tau\tau$ process where the two τ leptons further decay into an $e\mu$ pair with associated neutrinos. These neutrinos are originated in the subsequent decay process of the produced leptons from Z/γ^* boson. Therefore, they most likely carry less energies (implying lower $E_{\mathrm{T}}^{\mathrm{miss}}$ in the event) in comparison with the produced neutrinos in the direct decay of the two W bosons in a WW signal candidate. Hence, the Drell-Yan background can be suppressed by selecting events with large $E_{\mathrm{T}}^{\mathrm{miss}}$.

W+jets and Multi-jet

The leptonic decay of the W boson in the W+jets process produces one lepton and an associated neutrino. One of the hadronic jets in this process can be falsely reconstructed and identified as a lepton of opposite charge and different flavour to enter into the signal selection. In case of the multi-jet production process, two jets must be misidentified as leptons to mimic the WW signal. Since both of these processes contain jets that are misidentified as leptons, they are together considered as the fake background. A strict lepton selection - primarily identification and isolation criteria, significantly reduces this fake background contribution.

Other Diboson

The diboson processes, such as WZ, ZZ and $V+\gamma$ ($V = Z, W$) enter the analysis selection as minor backgrounds. For WZ and ZZ processes, the W boson decays into a lepton and a neutrino whereas the Z boson produces a lepton pair of opposite charge and same flavour. In case of the $V+\gamma$ process, γ produces an electron-positron pair by photon conversion and W or Z boson decays leptonically as aforementioned. The final state of each of these processes contains one e and one μ to enter the signal selection but is associated with one or more additional lepton(s). The requirement of exact one $e\mu$ pair with no additional leptons minimises these contributions.

5.2. Monte Carlo Samples

The pp collisions at the LHC often contain multiple high-p_{T} jets associated with the production of different processes. The modelling of these high jet multiplicities in the generated MC samples is either accomplished by the matrix element calculation or produced at the showering stage as described in Sections 2.2 and 2.3. This section discusses the different MC generator specifications used in this analysis to model the signal and background processes.

Pythia and Herwig++ are $2 \to 2$ generators that calculate the hard-scattered matrix element for two incoming and two outgoing particles. Both these generators are capable of showering that produces additional jets in the final state. Powheg is another common $2 \to 2$ matrix element generator with NLO accuracy in the perturbative QCD. It must be interfaced with a parton shower generator, such as Pythia or Herwig++ able to model the showering process. Sherpa and MadGraph are $2 \to n$ generators that create the jet multiplicities in both matrix element calculation and showering step. An ambiguity resolution is performed to choose any one of the calculations to create the additional jets at a fixed order prediction in α_s. MC generators use Parton Distribution Functions (PDFs) to model the proton structure in the collisions accurately.

The following section discusses the chosen PDF set and MC generators that calculate the matrix element and parton shower in order to model the signal and background processes.

Signal Modelling

The WW final state produced by the $q\bar{q}$ interaction is generated with PowhegBox v2 [74–77] generator at next-to-leading order (NLO) matrix element calculation in the perturbative QCD using the CT10 NLO PDFs [78]. The patron shower and non-perturbative effects are modelled with Pythia v8.210 [79] and CTEQ6L1 [80] PDF set. The $q\bar{q} \to WW$ signal cross section provided at NLO is normalised to the next-to-next-to-leading order (NNLO) predicted cross section [81].

The resonant $gg \to H \to WW$ signal contribution is generated similarly with PowhegBox v2 [82] interfaced with Pythia v8.210 and normalised to the inclusive next-to-next-to-next-to-leading order (N^3LO) predicted cross section [83].

The non-resonant $gg \to WW$ signal contribution is modelled with Sherpa v2.1.1 at leading order (LO) using OpenLoops with up to one additional parton in the final state [84] and normalised using the inclusive NLO predicted cross section [85].

Background Modelling

The $t\bar{t}$ and Wt background processes are simulated with the PowhegBox v2 [86, 87] event generator and CT10 NLO PDFs. Parton shower and non-perturbative effects are simulated with Pythia v6.428 [88] and CTEQ6L1 PDF set. The top quark mass is explicitly set to 172.5 GeV in the simulation. The $t\bar{t}$ cross section is normalised to the NNLO+NNLL soft gluon resummation prediction [89] while the Wt sample is normalised to the NLO+NNLL prediction [90].

The Drell-Yan ($Z/\gamma^* \ (\to ee/\mu\mu/\tau\tau)$ + additional jets) production processes are simulated with the MadGraph5_aMC@NLO v2.2.2 [91] event generator interfaced to Pythia v8.186. The matrix elements for the Z production with up to four associated partons are calculated at LO. It uses the NNPDF2.3 LO PDF set [92]. The PHOTOS++ program at version 3.52 [93] is used for QED emissions from electroweak final state particles. The W+jets process is generated with PowhegBox v2 generator interfaced with Pythia v8.210. Both Drell-Yan and W+jets cross sections are normalised to NNLO [94] predictions.

The WZ and ZZ processes are simulated using the PowhegBox v2 generator similar to the $q\bar{q} \to WW$ process where the parton shower and non-perturbative effects are modelled with Pythia v8.210. The cross sections are normalised to NNLO [95, 96] predictions.

The Sherpa v2.1.1 event generator is used to model the $V + \gamma$ ($V = Z, W$) processes with LO matrix element calculations. The generator produces events with up to three partons in the final state in both matrix element calculation and parton shower. The CT10 NLO PDF set is

used and the transverse momentum of the photon γ is set to be greater than 10 GeV in the simulation.

The EvtGen v1.2.0 [97] program is used for the properties of the hadron (primarily containing b and c quarks) decays in all samples generated using the PowhegBox v2 and MadGraph5_aMC@NLO v2.2.2 programs.

The nominal samples to model the WW signal and background processes and their cross section normalisations to the highest order perturbative predictions are summarised in Table 5.1. The MC samples are expressed by the name of the used generators omitting the version specifications throughout this thesis unless explicitly stated.

Process	Generator	Normalisation
	Signal Processes	
$qq \to WW$	Powheg+Pythia8	NNLO
$gg \to H \to WW$	Powheg+Pythia8	N^3LO
$gg \to WW$	Sherpa 2.1	NLO
	Background Processes	
$t\bar{t}$	Powheg+Pythia6	NNLO+NNLL
Wt	Powheg+Pythia6	NLO+NNLL
Drell-Yan (Z + jets)	MadGraph+Pythia8	NNLO
WZ	Powheg+Pythia8	NNLO
ZZ	Powheg+Pythia8	NNLO
$V + \gamma$	Sherpa 2.1	LO
W+jets	Powheg+Pythia8	NNLO

Table 5.1.: Nominal MC samples used in the analysis for the modelling of the signal and background processes.

5.3. Event Selection

The WW event selection uses a set of selection cuts to be imposed on the kinematic distributions of the final state objects. These cut choices are meant to be applied to enhance the signal contribution over the background contaminations. The objects are required to satisfy the analysis specific trigger criteria before entering the final event selection. The analysis trigger and kinematic requirements imposed on the final state objects in the SR selection are detailed in the following.

Analysis Trigger Selection

Data events must satisfy both the hardware-based L1 and software-based HLT trigger criteria to be recorded by the ATLAS TDAQ system (Subsection 3.2.4). Each analysis further requires a particular choice of analysis specific trigger requirements to be satisfied by the selected objects in an event. Both data and MC have these analysis triggers implemented so that a proper data to MC treatment can be pursued. The WW analysis applies individual trigger selections on the final state leptons. The chosen single lepton triggers and the assigned L1 and HLT trigger p_{T} thresholds are listed in Table 5.2. Data and MC events enter the analysis selection if at least one of the leptons in the event satisfies the assigned single lepton trigger requirement. Events satisfying other trigger requirements than the ones listed are considered non-interesting for this analysis. The lepton trigger efficiencies in data and MC to pass the analysis trigger selection is used to calculate the trigger SFs. The SFs are then applied to the MC to replicate the data trigger efficiency as mentioned in Section 4.8. The trigger SFs also compensate the difference in L1 electron trigger p_{T} threshold between data and MC as stated in Table 5.2.

Trigger Name	L1 p_{T} (GeV)	HLT p_{T} (GeV)
Single-e-Trigger		
e - `HLT_e24_lhmedium_L1EM18VH` OR `HLT_e60_lhmedium` (Data)	18	24 OR 60
e - `HLT_e24_lhmedium_L1EM20VH` OR `HLT_e60_lhmedium` (MC)	20	24 OR 60
Single-μ-Trigger		
μ - `HLT_mu20_iloose_L1MU15` OR `HLT_mu50`	15	20 OR 50
Analysis Trigger Selection		
Single-e-Trigger OR Single-μ-Trigger		

Table 5.2.: Analysis trigger selection imposed on the final state electron and muon candidates in an event.

Kinematic Requirements

The data and MC events selected by the analysis trigger criteria further satisfy the kinematic requirements in the SR as summarised in Table 5.3.

This analysis selects events containing an $e\mu$ pair with two high energetic ($p_{\mathrm{T}} > 25$ GeV) leptons of opposite charge and different flavour. Both leptons must satisfy identification and isolation criteria that reduce the fake (W+jets and multi-jet) background containing jets misidentified as leptons. The di-lepton invariant mass $m_{e\mu}$ of greater than 10 GeV rejects the phase space where the backgrounds are badly modelled in the MC samples. Events with additional leptons (with $p_{\mathrm{T}} > 10$ GeV) are not considered in this analysis to reduce the other diboson (WZ, ZZ and $W/Z + \gamma$) backgrounds. The $t\bar{t}$ and Wt processes contain jets (mainly b-jets) in their final states in addition to the $e\mu$ pair. Events containing jets with p_{T} >25

Object requirements for final event selection	
Lepton selection	
Lepton pair selection	$e\mu$ pair, opposite charge
Lepton transverse momentum $p_{\mathrm{T}}^{e,\mu}$	> 25 GeV
Lepton identification	e - Tight, μ - Medium
Lepton isolation	Gradient isolation
Di-lepton invariant mass $m_{e\mu}$	> 10 GeV
Additional background specific selection	
Number of additional leptons N_{add}^{ℓ}	$= 0$
Number of jets N_{jets}	$= 0$
Number of b-jets $N_{\text{b-jets}}$	$= 0$
$E_{\mathrm{T,rel}}^{\mathrm{miss}}$	> 15 GeV
$p_{\mathrm{T}}^{\mathrm{miss}}$	> 20 GeV

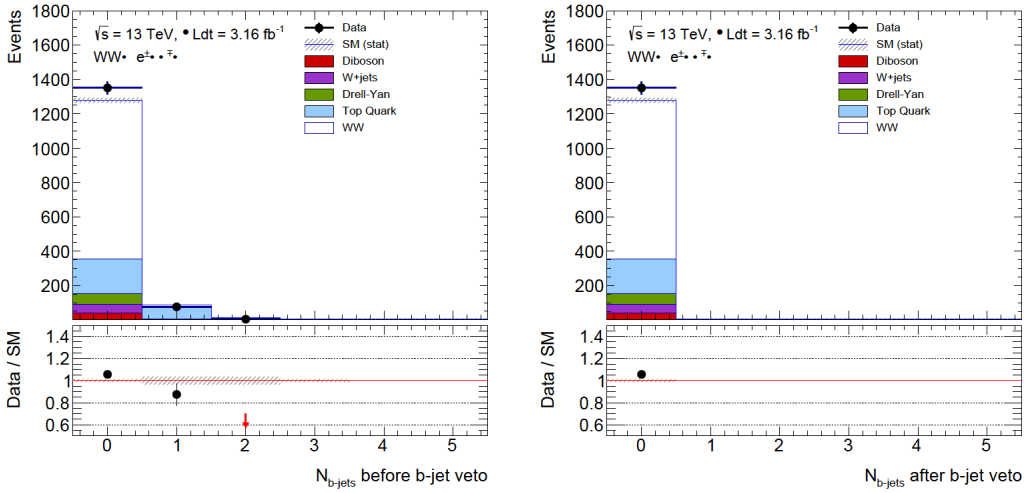
Table 5.3.: Event selection criteria to select WW signal events in the SR.

(30) GeV in the region $|\eta| < 2.4$ (4.5) are removed from the analysis selection to minimise the top contribution. An additional requirement of discarding events containing b-jets ($p_{\mathrm{T}} > 20$ GeV, $|\eta| < 2.4$) is incorporated due to the jet p_{T} threshold difference in jet and b-jet selections. This requirement rejects b-jets with p_{T} between 20 to 25 GeV and further suppresses the top background. Finally, the Drell-Yan background is reduced by requiring higher $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ and $p_{\mathrm{T}}^{\mathrm{miss}}$ in the event coming from the final state neutrinos. For $Z/\gamma^* \to \tau\tau$ events where the $E_{\mathrm{T}}^{\mathrm{miss}}$ tends to align with the final state leptons, the $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ variable becomes more efficient than $E_{\mathrm{T}}^{\mathrm{miss}}$ in reducing this background contribution.

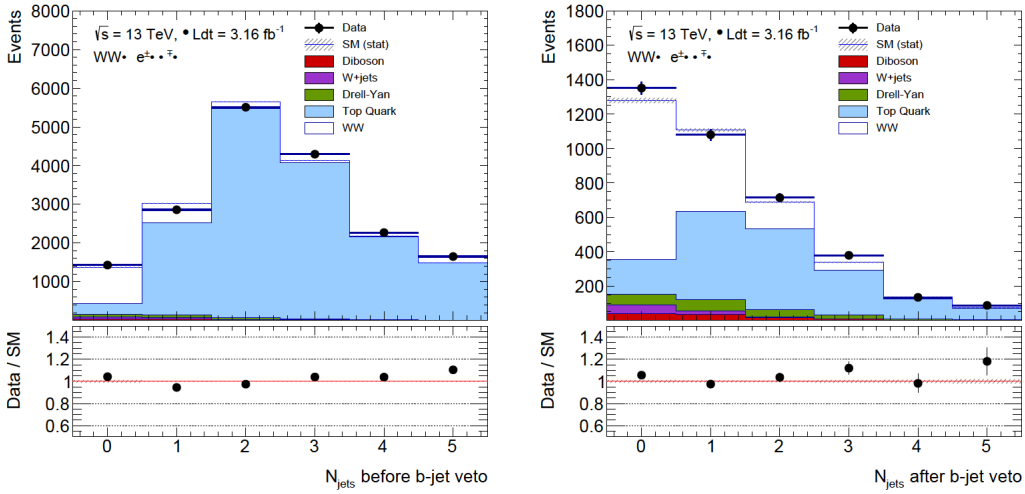
The effectiveness of the b-jet veto ($N_{\text{b-jets}} = 0$) selection in reducing the top contribution is demonstrated in both Figures 5.1a and 5.1b. In Figure 5.1a, the b-jet multiplicity distribution is shown for events fulfilling all selection cuts listed in Table 5.3 but with or without the b-jet veto cut. Events with one or more b-jets ($20 < p_{\mathrm{T}} < 25$ GeV) after the jet veto ($N_{\mathrm{jets}} = 0$) selection are rejected by this b-jet veto requirement. The jet multiplicity distribution is also presented for events satisfying all selection requirements except the jet veto and with or without the b-jet veto selection where the reduction of the top background fraction in the total selected events can also be visualised.

Table 5.4 compares the total number of the predicted signal and background events with the data satisfying each of the SR event selection cuts. For more straightforward demonstration, the estimated MC events are approximated to the nearest integer. In Table 5.5, the relative contributions of signal and background processes passing the different selection cuts are also shown. It illustrates the decrease in different backgrounds and increase in the signal fraction at different selection cut stages. The estimated event yields for the three signal subprocesses and their relative compositions in the total signal prediction are presented in Table 5.6.

The data and MC comparisons of the lepton kinematic and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions after the full event selection are shown in Figure 5.2.



(a) The b-jet multiplicity distributions before (left) and after (right) the b-jet veto selection with all other event selection cuts applied.



(b) The jet multiplicity distributions before (left) and after (right) the b-jet veto selection with all other event selection cuts except the jet veto cut applied.

Figure 5.1.: Data/MC comparisons of the b-jet and jet multiplicity distributions before (left) and after (right) the b-jet veto selection. The uncertainty only includes the statistical component coming from the data and MC sample statistics.

$e\mu + \mu e$ Channel	Top	Drell-Yan	WZ/ZZ	V/γ	W+jets	WW Signal	Data	Data/MC
Pre-selection	37768 ± 27	13049 ± 121	1107 ± 2	1168 ± 16	3362 ± 79	3994 ± 13	94985	1.94
Lepton Selection	22855 ± 21	2784 ± 56	311 ± 1	134 ± 5	136 ± 16	2407 ± 11	29041	1.66
N_{add}^{ℓ} =0	22715 ± 21	2731 ± 55	109 ± 1	109 ± 4	134 ± 16	2405 ± 11	28562	1.02
$N_{\text{jets}} = 0$	347 ± 3	1765 ± 50	29 ± 0	50 ± 3	70 ± 12	1278 ± 8	3753	1.08
$N_{\text{b-jets}} = 0$	247 ± 2	1745 ± 49	28 ± 0	50 ± 3	66 ± 11	1265 ± 8	3627	1.09
$E_{\text{T,rel}}^{\text{miss}} > 15$ GeV	211 ± 2	215 ± 17	21 ± 0	24 ± 2	52 ± 10	996 ± 7	1565	1.07
$p_{\text{T}}^{\text{miss}} > 20$ GeV	203 ± 2	60 ± 9	19 ± 0	18 ± 2	52 ± 10	924 ± 7	1351	1.10
Signal Region	203 ± 2	60 ± 9	19 ± 0	18 ± 2	52 ± 10	924 ± 7	1351	1.10

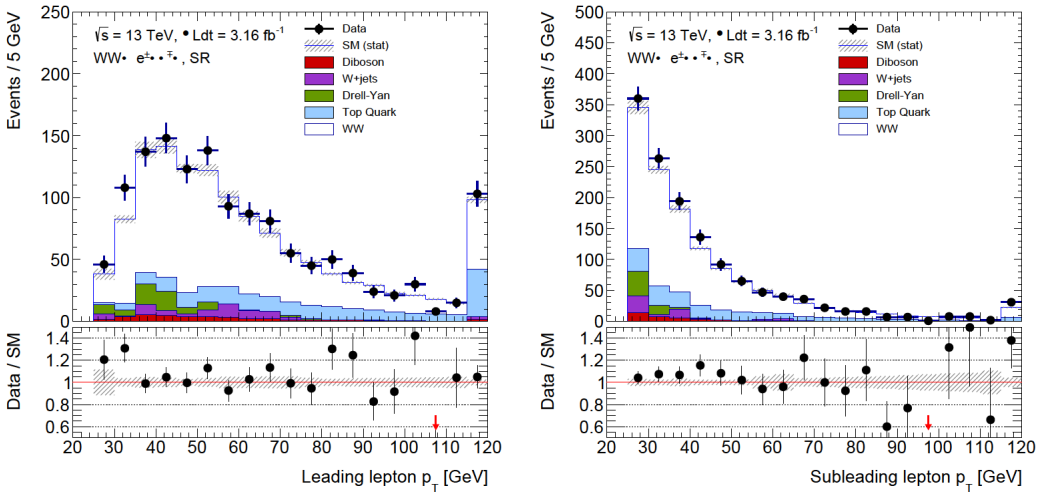
Table 5.4.: Comparisons of the signal and background predictions with the observed data at each event selection requirement. Only the statistical uncertainty originating from the sample statistics is shown.

$e\mu + \mu e$ Channel	Top	Drell-Yan	WZ/ZZ	V/γ	W+jets	WW Signal	S/B
Pre-selection	0.62	0.22	0.02	0.02	0.06	0.07	0.07
Lepton Selection	0.80	0.10	0.01	0.001	0.001	0.09	0.09
N_{add}^{ℓ} =0	0.81	0.10	0.01	0.001	0.001	0.09	0.09
$N_{\text{jets}} = 0$	0.10	0.50	0.01	0.01	0.02	0.36	0.57
$N_{\text{b-jets}} = 0$	0.07	0.51	0.01	0.01	0.02	0.37	0.59
$E_{\text{T,rel}}^{\text{miss}} > 15$ GeV	0.14	0.14	0.01	0.02	0.03	0.66	1.91
$p_{\text{T}}^{\text{miss}} > 20$ GeV	0.16	0.05	0.01	0.01	0.04	0.72	2.62
Signal Region	0.16	0.05	0.01	0.01	0.04	0.72	2.62

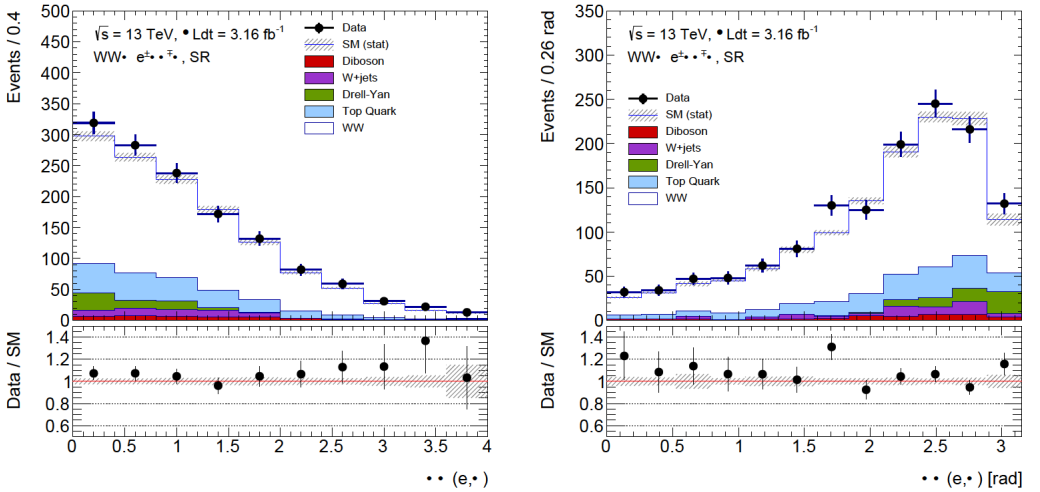
Table 5.5.: Fractional compositions of signal and background predictions at each event selection requirement.

$e\mu + \mu e$ Channel	WW Signal	$q\bar{q} \to WW$		$gg \to WW$		$gg \to H \to WW$	
		N_{events}	Frac.	N_{events}	Frac.	N_{events}	Frac.
Pre-selection	3994 ± 13	3553 ± 13	0.89	238 ± 3	0.06	203 ± 1	0.05
Lepton Selection	2407 ± 11	2207 ± 10	0.92	139 ± 2	0.06	62 ± 0	0.03
N_{add}^{ℓ} =0	2405 ± 11	2205 ± 10	0.92	138 ± 2	0.06	62 ± 0	0.03
$N_{\text{jets}} = 0$	1278 ± 8	1194 ± 8	0.93	66 ± 1	0.05	19 ± 0	0.01
$N_{\text{b-jets}} = 0$	1265 ± 8	1183 ± 8	0.93	64 ± 1	0.05	18 ± 0	0.01
$E_{\text{T,rel}}^{\text{miss}} > 15$ GeV	996 ± 7	921 ± 7	0.93	58 ± 1	0.06	17 ± 0	0.02
$p_{\text{T}}^{\text{miss}} > 20$ GeV	924 ± 7	852 ± 6	0.92	56 ± 1	0.06	16 ± 0	0.02
Signal Region	924 ± 7	852 ± 6	0.92	56 ± 1	0.06	16 ± 0	0.02

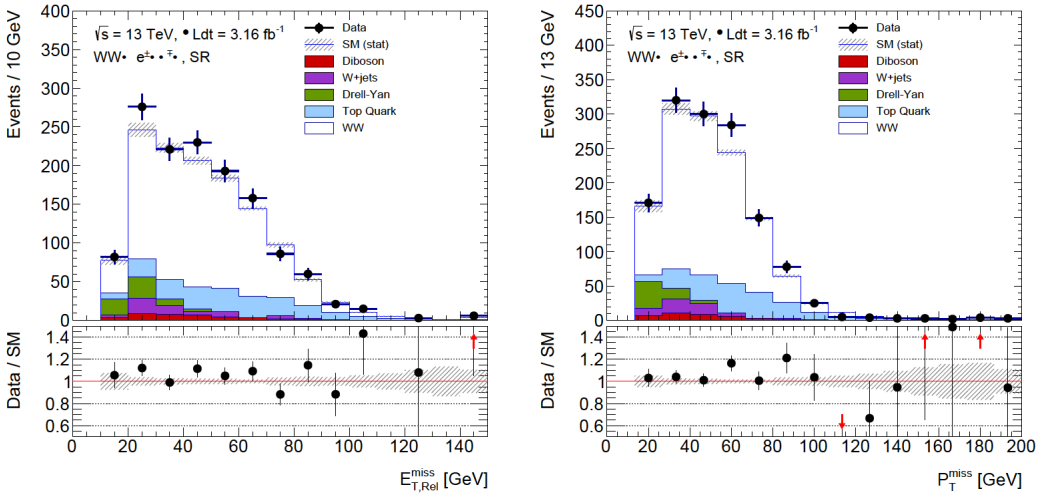
Table 5.6.: Event yields of the total signal prediction and the fractional contributions of the three subprocesses at different event selection requirements. Only the statistical uncertainty coming from the signal MC sample statistics is included.



(a) Transverse momentum p_{T} of the leading (left) and subleading (right) leptons.



(b) Differences in pseudorapidity $\Delta\eta$ (left) and azimuthal angle $\Delta\phi$ (right) between the leptons in the $e\mu$ pair.



(c) Relative missing transverse momentum $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ (left) and track-based missing transverse momentum $p_{\mathrm{T}}^{\mathrm{miss}}$ (right) associated with the neutrinos in the final state.

Figure 5.2.: Data/MC comparisons of the kinematic distributions of final state leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$ related quantities fullfilling the SR event selection. Only the statistical uncertainty is included.

5.4. Background Determination

After applying the event selection in the SR, the dominant background contribution comes from the top ($t\bar{t}$+Wt) process. The contribution of the top background is 16% of the total (signal+background) MC prediction as shown in Table 5.5. The Drell-Yan and fake (W+jets and multi-jet) processes are the second and third largest background components to enter the SR with contributions of 5% and 4% respectively. Other diboson processes (WZ, ZZ and $V+\gamma$) play a minor role with 2% contribution in the selected MC events in the SR.

The two most significant backgrounds in this analysis are the top and Drell-Yan processes. Both these backgrounds are estimated using a data-driven normalisation method called *Transfer Factor*. In this method, the MC predicted background is normalised to data in a dedicated Control Region (CR), enriched with that particular background process. The normalisation is then propagated to the SR to estimate the background contribution. The methodology and applicability of this approach are discussed in the relevant subsections (Subsection 5.4.1 and 5.4.2).

This normalisation method cannot be applied to processes with low statistics of the available MC samples or due to difficulties involved in isolating these processes from the signal. In such case, backgrounds are derived directly from data using specific methods. In this analysis, the fake background contribution is derived using the *Matrix Method* - a data-driven method that is described in Subsection 5.4.3. The *Matrix Method* utilises the efficiencies of real (leptons from W boson's decay in this case) and fake (jet misidentified as a lepton, leptons from semileptonic decays of heavy flavour jets) leptons to satisfy the SR selection and therefore estimates the fake background contribution in data.

In case of subdominant backgrounds, the MC-based predictions are considered for processes having sufficient MC sample statistics. Subsection 5.4.4 discusses the other diboson backgrounds (WZ, ZZ and $V+\gamma$) estimated using this MC-based predictions.

For the aforementioned background estimations, specific validation regions are considered to validate the MC modelling and involved methodologies.

5.4.1. Top Background

The top ($t\bar{t}$+Wt) background contribution in the SR is highly suppressed by removing events containing any jets or b-jets. The jet algorithms select jets with $p_{\mathrm{T}} > 25$ (30) GeV for $|\eta| < 2.4$ (4.5) whereas the b-tagging algorithm identifies jets coming from b quarks with $p_{\mathrm{T}} > 20$ GeV in the central region of the detector ($|\eta| < 2.4$) (Section 4.4).

The remaining top contribution in the SR (16% of the total MC prediction) originates from events with jets either falling below the jet p_{T} threshold or being non-identified b-jets by the

b-tagging algorithm. The top background is estimated using the *Transfer Factor* method as described below.

Transfer Factor Method

In this approach, a specific Control Region (CR) is constructed in which the dominant contribution comes from a particular background process of interest. The MC prediction of the background is normalised to data in that CR. The normalisation factor is then applied to the SR MC prediction of that background process. This results in the final estimation of that background contribution in the SR.

The CR is constructed in such a way that the background normalisation obtained in the CR should apply to the SR as well. Therefore, the CR maintains minimal extrapolation and orthogonality to the SR. To achieve that, the CR closely follows the SR selection but explicitly reverses/fails one or few SR cuts. The cuts are reversed to enhance the contribution of that particular background in the CR. This cut reversal also causes the CR to be orthogonal to the SR. The orthogonality requirement makes the CR statistically independent to the SR. Therefore, in the constructed CR, the signal contribution is expected to be negligible and the obtained normalisation factor should not be affected by the signal.

Top Control Region

For the top background estimation using this approach, the normalisation is performed in a top control region, denoted as the TopCR. The TopCR uses the following selections:

- Identical lepton selection as in the SR selection (Table 5.3).
- Number of additional leptons $N^{\ell}_{\text{add}} = 0$.
- Invariant mass of the $e\mu$ pair $m_{e\mu} > 10$ GeV.
- $p_{\text{T}}^{\text{miss}} > 20$ GeV but no $E_{\text{T,rel}}^{\text{miss}}$ cut. This is because of the sensitivity of $E_{\text{T,rel}}^{\text{miss}}$ variable with the nearby jets.
- Exactly one jet ($N_{\text{jets}} = 1$) and at least one b-jet ($N_{\text{b-jets}} \geq 1$) in an event. Since the p_{T} thresholds are different for jet and b-jet selections, events with two or more b-jets would also be accepted in this region. This can only happen for b-jets with p_{T} between 20 and 25 GeV.

Requiring one jet in the event accomplishes the orthogonality requirement of the TopCR to the SR and provides a high top purity[7] phase space. The additional b-jets further enhance the fraction of top events in the TopCR.

[7] Purity is determined as the fraction of events coming from a particular process to the total number of events.

The top background is estimated from the normalisation in the TopCR using the following Equation 5.1.

$$N_{top}^{SR} = N_{top,MC}^{SR} \cdot \left(\frac{N_{data}^{CR} - N_{non-top,MC}^{CR}}{N_{top,MC}^{CR}} \right) \tag{5.1}$$

where N_{top}^{SR} is the number of top events in the SR estimated using the *Transfer Factor* method. $N_{top,MC}^{SR}$ represents the predicted top events in the SR by the nominal POWHEG+PYTHIA6 MC sample. N_{data}^{CR}, $N_{top,MC}^{CR}$ and $N_{non-top,MC}^{CR}$ are data and MC predicted number of top and non-top events in the TopCR. The non-top contribution involves physics processes excluding the top process. The bracketed term determines the normalisation factor obtained in the TopCR.

Equation 5.1 can be further rewritten as

$$N_{top}^{SR} = TF_{top} \cdot (N_{data}^{CR} - N_{non-top,MC}^{CR}), \quad TF_{top} = \frac{N_{top,MC}^{SR}}{N_{top,MC}^{CR}} \tag{5.2}$$

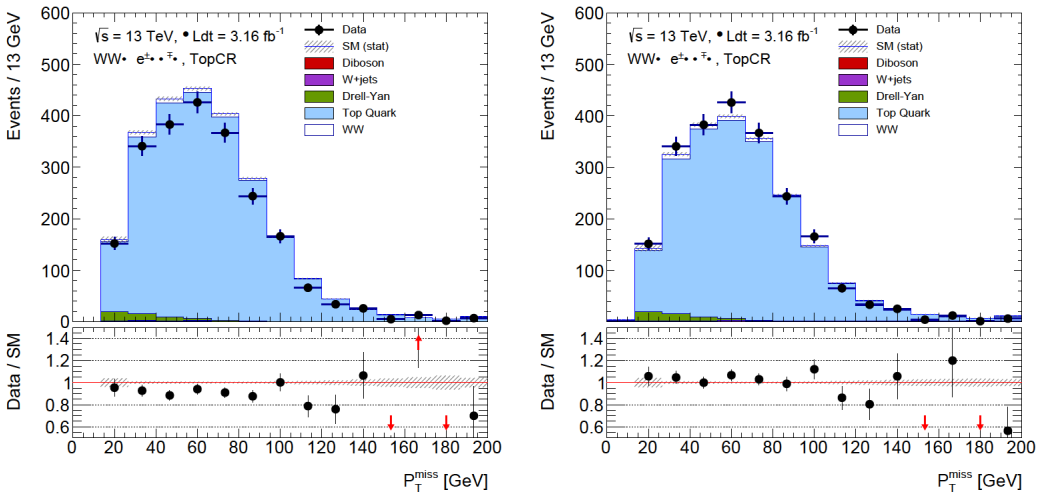
where the TF_{top} propagates the data-driven top estimate $(N_{data}^{CR} - N_{non-top,MC}^{CR})$ in the TopCR into the SR. The event yields of the top background and data are shown in Table 5.7 for both the TopCR and SR phase space. The top MC predictions summarised in this table are before applying the normalisation factor obtained in the TopCR. The purity of the top process in the TopCR is 93% and the non-top events mainly come from the Drell-Yan and WW signal processes with individual contributions of 2%. The impact of the WW signal contribution on the calculated TF_{top} or normalisation factor in the TopCR is appropriately treated in the final cross section calculation as will be discussed in the next chapter. An 11% data/MC discrepancy is observed in the TopCR. Hence, the data/MC compatibility is further checked in a validation region as described in the following to understand the MC modelling and also the applicability of the TopCR normalisation.

$e\mu + \mu e$ Channel	$t\bar{t}$	Wt	$t\bar{t}$+Wt	Top Purity	Total MC	Data	Data/MC
SR	129.5 ± 1.6	73.5 ± 1.1	203.1 ± 1.9	0.16	1300.4 ± 22.0	1351	1.04 ± 0.03
TopCR	1858.0 ± 6.0	493.1 ± 2.7	2351.2 ± 6.6	0.93	2520.0 ± 27.9	2232	0.89 ± 0.02

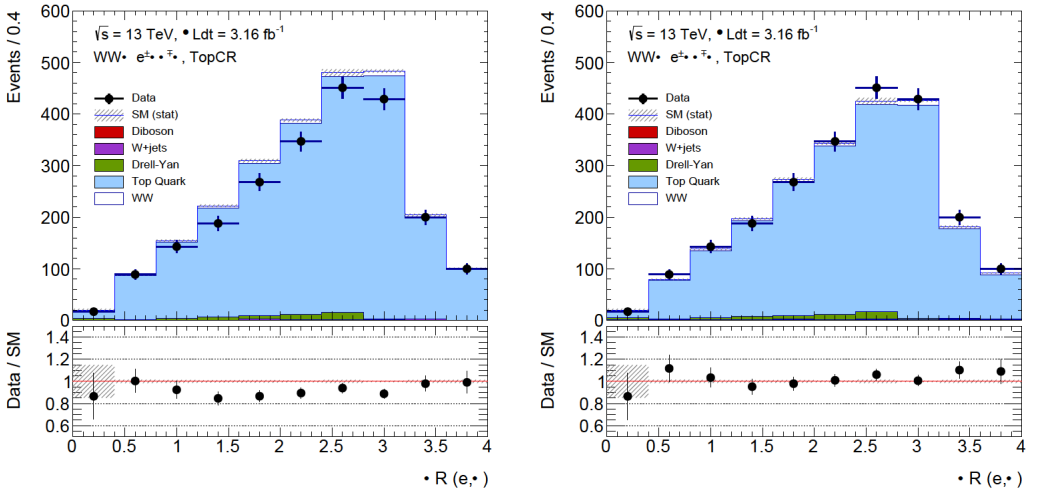
Table 5.7.: Yields comparison in SR and TopCR without applying the normalisation factor to the top MC prediction. Only the statistical uncertainty is included.

The data/MC comparisons of the kinematic distributions of E_T^{miss} and lepton related quantities in the TopCR are shown in Figure 5.3. These distributions are presented before and after applying the normalisation factor of 0.8775 (calculated using the bracketed term of Equation 5.1) to the top MC prediction. This is to demonstrate the applicability of the top background

normalisation obtained using the *Transfer Factor* method. The uncertainty on the normalisation factor has not been incorporated in these distributions.



(a) Track-based missing transverse momentum $p_{\mathrm{T}}^{\mathrm{miss}}$ before (left) and after (right) applying the normalisation factor to the top MC prediction.



(b) ΔR between the two leptons in the $e\mu$ pair before (left) and after (right) applying the normalisation factor to the top MC prediction.

Figure 5.3.: Data/MC comparisons of the kinematic distributions related to $E_{\mathrm{T}}^{\mathrm{miss}}$ and $e\mu$ system in the TopCR before (left) and after (right) the normalisation factor being applied to the top MC prediction. Only the statistical uncertainty is included.

Validation Region

The TopCR accepts events containing one jet and at least one b-jet in the final state while the SR requires no jet activity. Therefore, the extrapolation of TopCR to the SR is large. The validity of the used MC modelling is checked in another top dominated phase space close to the SR. The Top Validation Region (TopVR) is formed following the same selections as the CR

but with different jet and b-jet requirements. Events with no jets with $p_{\mathrm{T}} > 25$ GeV and at least one b-jet with $20 < p_{\mathrm{T}} < 25$ GeV are allowed in the TopVR. In Table 5.8, event yields are shown for both TopCR and TopVR without normalising the top MC predictions. The data/MC comparison of a di-lepton kinematic quantity in the TopVR is presented in Figure 5.4. The same distribution is presented before and after scaling the top MC prediction by the obtained normalisation factor in the TopCR. The uncertainties on the normalisation factor are not being considered in the scaled distribution.

$e\mu + \mu e$ Channel	$t\bar{t}$+Wt	Non Top	Top Purity	Total MC	Data	Data/MC
TopCR $N_{\text{jets}} = 1,\ N_{\text{b-jets}} \geq 1$	2351.2 ± 6.6	166.2 ± 27.1	0.93	2520.0 ± 27.9	2232	0.89 ± 0.02
TopVR $N_{\text{jets}} = 0,\ N_{\text{b-jets}} \geq 1$	91.2 ± 1.3	7.9 ± 3.4	0.92	104.4 ± 3.7	96	0.92 ± 0.10

Table 5.8.: Yields comparison in TopCR and TopVR without applying the normalisation factor to the top MC prediction. Only the statistical uncertainty is included.

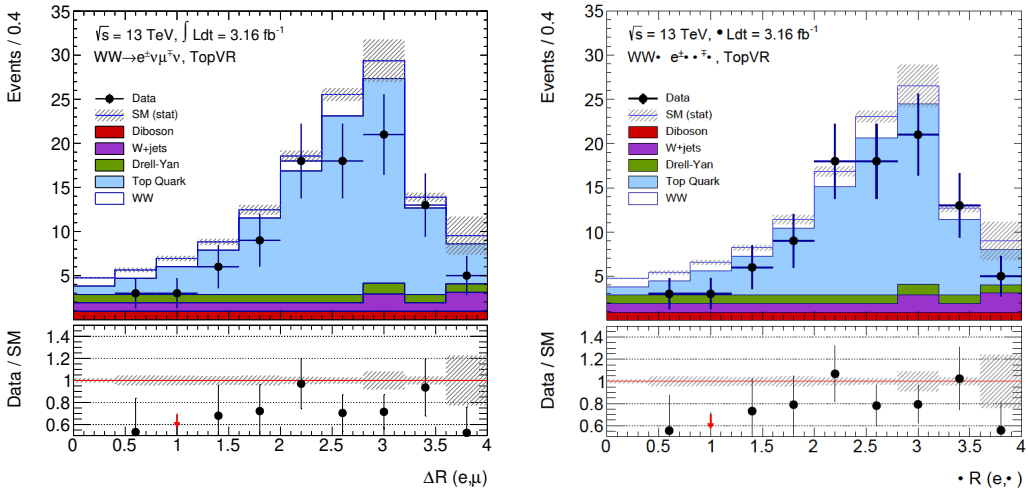


Figure 5.4.: Data/MC comparison of ΔR distribution in the TopVR before (left) and after (after) applying the normalisation factor to the top MC prediction. Only the statistical uncertainty is included.

Both of these regions (TopCR and TopVR) offer equivalent top purity and show a similar level of data/MC difference. However, it is worth to mention that the validation region is statistically limited. The normalisation obtained in the TopCR can be well applicable to both these phase spaces. The choice of the nominal top MC is further validated in two more validation regions, discussed in Appendix A.1.

Uncertainties on the Top Estimation

The top background is estimated using Equation 5.2 where the top estimation in the TopCR is propagated by the measured TF_{top} to the SR.

The statistical uncertainty associated with this estimate originates from the term ($N^{CR}_{data} - N^{CR}_{non-top}$) in the equation mentioned above. Therefore, the main statistical uncertainty comes from the data statistics in the TopCR. The statistical uncertainty from 7% contribution of the other non-top processes in the TopCR is negligible.

The systematic uncertainty on the top background is given by the total uncertainty associated with the transfer factor measurement. The transfer factor is measured from the nominal top MC sample and assigned a statistical uncertainty of 0.9% originating from the top MC sample statistics. The associated systematic uncertainty considers both the MC detector systematic uncertainties as well as uncertainties from MC modelling as discussed below.

Detector Uncertainties on TF_{top}

The effect of the MC detector systematic uncertainties (Section 4.8.2) on the TF_{top} is calculated. The dominant systematic uncertainties originate from the jet energy scale and b-tagging SF measurements. This is expected as the TF_{top} connects two regions with different jet activities.

Modelling Uncertainties on TF_{top}

The MC-based background estimations are affected by the theoretical modelling of the used nominal MC generators. The choice of a particular MC may bias the predictions. Therefore, the comparison of the nominal MC to alternative generators are assigned as the background modelling uncertainties that minimise the theoretical model dependence of the prediction. These uncertainties also incorporate physics effects not covered by the nominal MC due to the limitations of the model specific calculations behind it. The involved modelling uncertainties associated with the top estimation are discussed below.

- **Parton Shower (PS)** - The difference of the nominal Powheg+Pythia6 and a sample generated with Powheg matrix element calculation and Herwig++ parton shower modelling is considered.
- **Matrix Element (ME) Calculation** - The difference of Powheg+Herwig++ with aMC@NLO interfaced with Herwig++ is taken into account.
- **Initial State Radiation (ISR)** - The emission of gluons from the initial state partons further produce additional partons. This is denoted as the Initial State Radiation (ISR). For $t\bar{t}$ process, a separate sample is generated with Powheg+Pythia6 by varying the

parameters that affect extra jet radiations in both matrix element and parton shower calculations. The difference to the nominal MC is added as the ISR systematic uncertainty.

- **Diagram Removal and Diagram Subtraction (DS/DR)** - The feynman diagrams of the Wt process beyond the leading order calculations enter the $t\bar{t}$ process starting from the leading order. The diagram removal and subtraction approach remove the $t\bar{t}$ contribution from the Wt process at both amplitude and cross section level to avoid the interference between these two top processes. A comparison between the nominal Wt MC and an alternative sample generated with POWHEG+PYTHIA6 but implementing the diagram removal and subtraction approach is considered.
- ***tt̄* and *Wt* Cross Section** - The respective uncertainties associated with the $t\bar{t}$ and Wt cross sections of 6% and 10% are added as additional sources. These additional uncertainties cover the relative normalisations of the $t\bar{t}$ and Wt processes while calculating the transfer factor.

The full breakdown of statistical and systematic uncertainties on the top background estimation in the SR and TopCR and also on the $\mathrm{TF_{top}}$ are summarised in Table A.2. The $\mathrm{TF_{top}}$ is measured within 14.8% total uncertainty and the final value is given as

$$\mathrm{TF_{top}} = 0.0864 \pm 0.0128.$$

The estimated top background with associated statistical and systematic uncertainties in the SR using the *Transfer Factor* method is

$$N_{\mathrm{top}}^{\mathrm{SR}} = 178 \ \pm 4\,(\mathrm{stat.}) \ \pm 28\,(\mathrm{syst.}).$$

5.4.2. Drell-Yan Background

The Drell-Yan process ($Z/\gamma^*(\rightarrow ee/\mu\mu/\tau\tau)$+jets) is the second largest background (5% of the total MC prediction) to enter into the WW SR selection. The analysis requirement of selecting two opposite flavour leptons suppresses the contribution from $\mathrm{Z}/\gamma^* \rightarrow ee/\mu\mu$ process. The residual component comes mainly from $\mathrm{Z}/\gamma^* \rightarrow \tau\tau$ where the two τ leptons decay into an $e\mu$ pair with $E_{\mathrm{T}}^{\mathrm{miss}}$ from associated neutrinos. The $E_{\mathrm{T,rel}}^{\mathrm{miss}} > 15$ GeV and $p_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV selections further reduce the contamination due to this background in the SR.

The *Transfer Factor* approach as employed to the top background estimation is adapted to determine the Drell-Yan contribution in the SR. The background is normalised to data according to the following Equation 5.3 in a Drell-Yan control region, referred to as the DYCR. The normalisation obtained in the DYCR is extrapolated back to the SR by applying the transfer factor $\mathrm{TF_{DY}}$.

$$\mathrm{N}_{\mathrm{DY}}^{\mathrm{SR}} = \frac{\mathrm{N}_{\mathrm{DY,MC}}^{\mathrm{SR}}}{\mathrm{N}_{\mathrm{DY,MC}}^{\mathrm{CR}}} \cdot (\mathrm{N}_{\mathrm{data}}^{\mathrm{CR}} - \mathrm{N}_{\mathrm{non-DY,MC}}^{\mathrm{CR}}) = \mathrm{TF}_{\mathrm{DY}} \cdot (\mathrm{N}_{\mathrm{data}}^{\mathrm{CR}} - \mathrm{N}_{\mathrm{non-DY,MC}}^{\mathrm{CR}}) \tag{5.3}$$

where $\mathrm{N}_{\mathrm{DY}}^{\mathrm{SR}}$ is the estimated number of Drell-Yan events in the SR phase space using *Transfer Factor* method. $\mathrm{N}_{\mathrm{DY,MC}}^{\mathrm{SR}}$ represents the predicted Drell-Yan events in the SR by the nominal MadGraph+Pythia8 MC sample. $\mathrm{N}_{\mathrm{data}}^{\mathrm{CR}}$, $\mathrm{N}_{\mathrm{DY,MC}}^{\mathrm{CR}}$ and $\mathrm{N}_{\mathrm{non-DY,MC}}^{\mathrm{CR}}$ are the data and MC predicted number of Drell-Yan and non Drell-Yan events in the DYCR respectively.

Control Region

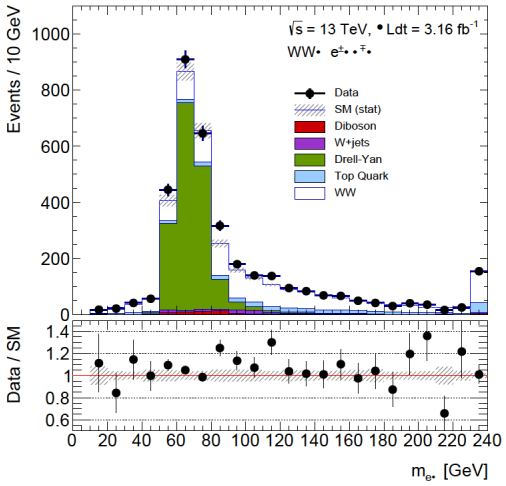
The DYCR uses the following selections:

- Identical lepton selection as applied in the SR event selection (Table 5.3).
- Number of additional leptons $N_{\mathrm{add}}^{\ell} = 0$.
- Invariant mass of the $e\mu$ pair $m_{e\mu} > 10$ GeV.
- No associated jets or b-jets in the event ($N_{\mathrm{jets}} = 0$ and $N_{\mathrm{b\text{-}jets}} = 0$).
- Selected Mass window on the invariant mass of the $e\mu$ pair $m_{e\mu} \in [45, 80]$ GeV.
- $E_{\mathrm{T,rel}}^{\mathrm{miss}} < 15$ GeV OR $p_{\mathrm{T}}^{\mathrm{miss}} < 20$ GeV.

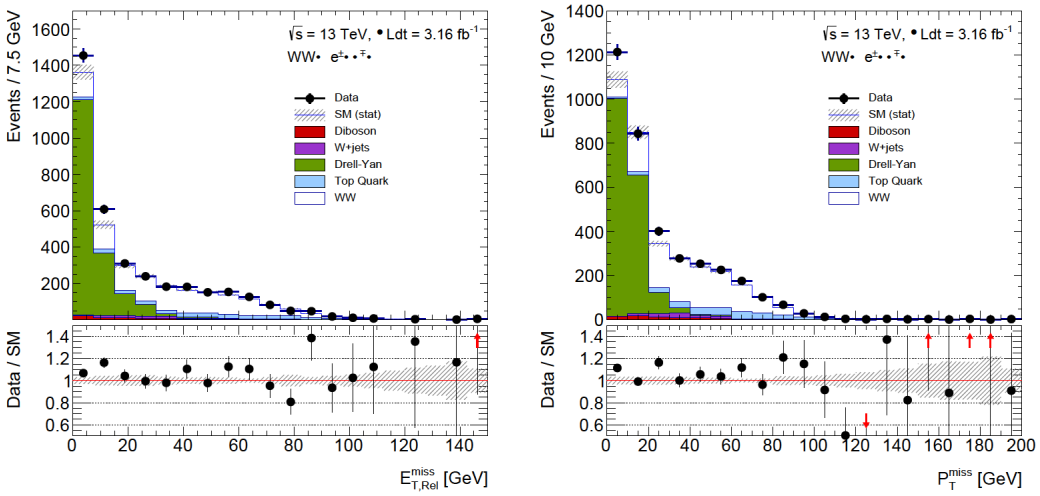
The invariant mass distribution of the di-lepton pair (mainly $\tau\tau$) from the Z boson decay is expected to peak around the Z boson invariant mass of $\sim$91 GeV [3]. The two τ leptons further produce an $e\mu$ pair with associated neutrinos in the subsequent decay process. Hence, the peak of the invariant mass distribution of the $e\mu$ pair is shifted towards a lower value because of the energy carried by the associated neutrinos. The mass window selection ($m_{e\mu} \in [45, 80]$ GeV) is motivated to enhance the Drell-Yan contribution in the DYCR as demonstrated in Figure 5.5a. The major part of the Drell-Yan events falls within this mass window range [45, 80] GeV.

The orthogonality of the control region is achieved by reversing the two $E_{\mathrm{T}}^{\mathrm{miss}}$ related criteria used in the SR. The $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ and $p_{\mathrm{T}}^{\mathrm{miss}}$ distributions shown in Figure 5.5b justify the $E_{\mathrm{T}}^{\mathrm{miss}}$ selections used to enhance the Drell-Yan events in the chosen control region.

The control region offers 95% Drell-Yan purity as listed in Table 5.9 with sufficiently high data and MC statistics. The events from processes other than the Drell-Yan mainly come from the WW signal with 4% contribution in the DYCR. The effect of this signal contribution on the calculated normalisation or transfer factor is correctly treated in the final cross section measurement as discussed in the next chapter. The high sample statistics in the DYCR significantly reduces any statistical fluctuations that may influence the estimation. The data/MC



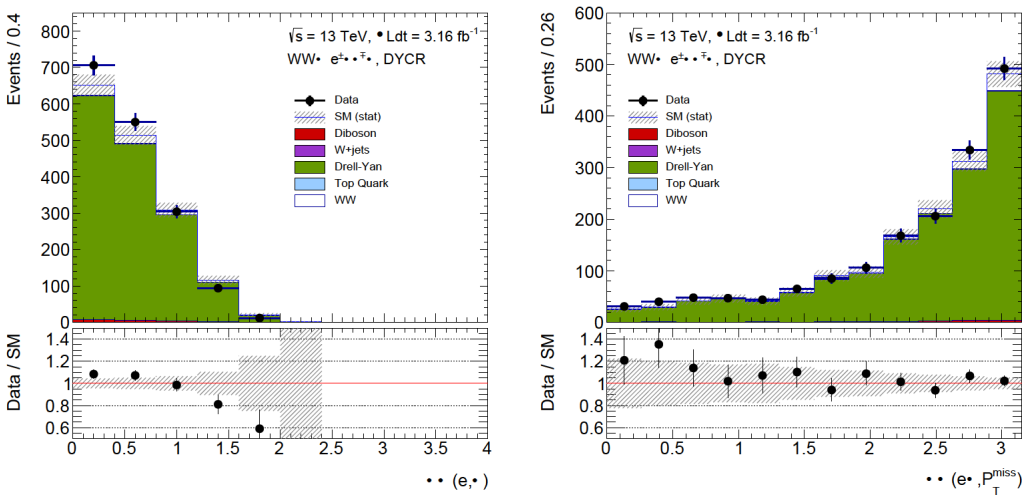
(a) Invariant mass of the $e\mu$ pair to motivate the mass window selection $m_{e\mu} \in [45, 80]$ GeV in the DYCR.



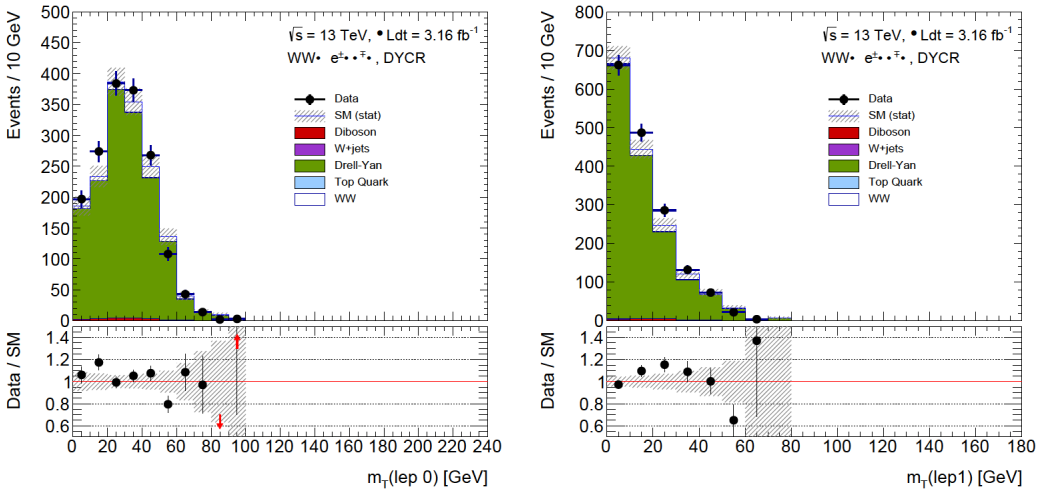
(b) $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ and $p_{\mathrm{T}}^{\mathrm{miss}}$ distributions to motivate the $E_{\mathrm{T,rel}}^{\mathrm{miss}} < 15$ GeV and $p_{\mathrm{T}}^{\mathrm{miss}} < 20$ GeV selections used in the DYCR.

Figure 5.5.: Data/MC comparisons of $m_{e\mu}$ and $E_{\mathrm{T}}^{\mathrm{miss}}$ related distributions to motivate the chosen DYCR selection. The distributions shown here are after applying the lepton selection, no additional leptons, jet veto and b-jet veto selections. The uncertainty covers only the statistical component coming from the data and MC sample statistics.

comparisons of the lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ related quantities in the DYCR are presented in Figure 5.6 without applying the normalisation factor obtained in the DYCR to the Drell-Yan MC prediction.



(a) $\Delta\eta$ between the two leptons in the $e\mu$ pair (left) and $\Delta\phi$ between the $e\mu$ system and $E_{\mathrm{T}}^{\mathrm{miss}}$ object.



(b) Transverse mass of the leading (left) and subleading (right) leptons. The transverse mass of a lepton is defined as $\mathrm{m_T(lep\ x)} = \sqrt{2p_{\mathrm{T}}^{\ell} \cdot p_{\mathrm{T}}^{\mathrm{miss}} \cdot (1 - \cos\Delta\phi)}$ where $x = 0$ and 1 represent the leading and subleading leptons respectively.

Figure 5.6.: Data/MC comparisons of lepton related quantities in the DYCR. The normalisation factor has not been applied to the Drell-Yan MC prediction. Only the statistical uncertainty is included.

Validation Region

The Drell-Yan background is validated in another potential orthogonal phase space close to the SR. The Drell-Yan Validation Region (DYVR) is constructed using the same lepton, jet and

mass window selections as the DYCR but only altering the $E_{\mathrm{T}}^{\mathrm{miss}}$ selection combination. The $E_{\mathrm{T}}^{\mathrm{miss}}$ criteria used in the DYVR is

$$(E_{\mathrm{T,rel}}^{\mathrm{miss}} < 15 \text{ GeV and } p_{\mathrm{T}}^{\mathrm{miss}} > 20 \text{ GeV}) \text{ OR } (E_{\mathrm{T,rel}}^{\mathrm{miss}} > 15 \text{ GeV and } p_{\mathrm{T}}^{\mathrm{miss}} < 20 \text{ GeV}).$$

This region also provides a Drell-Yan purity of $\sim 77\%$ but with insufficient statistical power as summarised in Table 5.9. The data/MC comparisons for lepton related distributions are shown in Figure 5.7 for this validation region. The Drell-Yan MC prediction is not normalised by the normalisation factor in these shown distributions. Both DYCR and DYVR show a good data/MC agreement within the associated statistical uncertainties.

$e\mu + \mu e$ Channel	Drell-Yan	$ee/\mu\mu$ Fraction	DY Purity	Total MC	Data	Data/MC
SR	60.4 ± 8.6	0.08 ± 0.05	0.05	1300.3 ± 22.0	1351	1.04 ± 0.03
DYCR	1515.7 ± 45.8	0.02 ± 0.01	0.95	1596.6 ± 48.4	1666	1.04 ± 0.04
DYVR	168.6 ± 15.3	0.07 ± 0.03	0.77	218.7 ± 16.9	205	0.94 ± 0.10

Table 5.9.: Yields comparison in the SR, DYCR and DYVR. The $ee/\mu\mu$ fraction denotes the percentage contribution of $Z/\gamma^* \to ee/\mu\mu$ process in the total Drell-Yan background. The Drell-Yan events reported in this table are before the normalisation factor being applied. Only the statistical uncertainty is included.

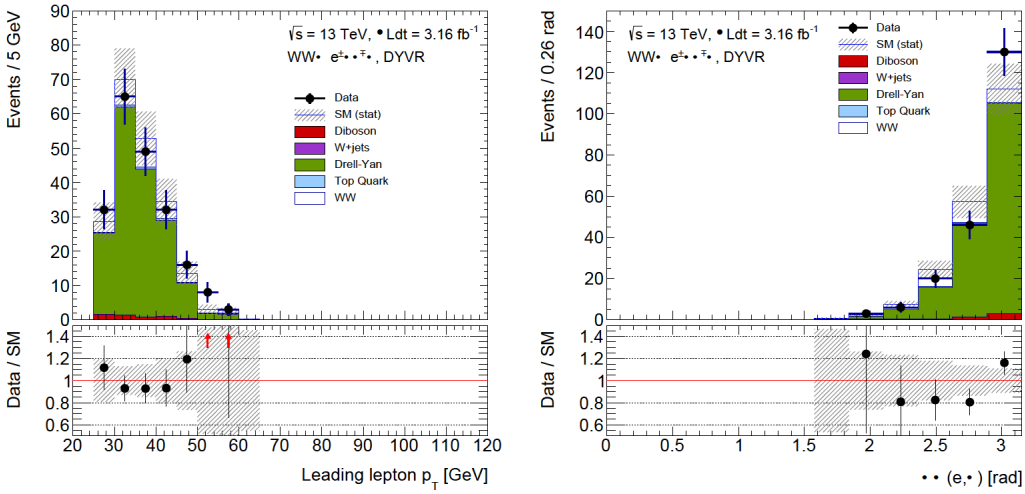


Figure 5.7.: Data/MC comparisons of the kinematic distributions - the leading lepton p_{T} (left) and $\Delta\phi$ between the two leptons in the $e\mu$ pair (right) in the DYVR. The normalisation factor has not been applied to the Drell-Yan MC prediction. The uncertainty covers only the statistical component coming from the data and MC sample statistics.

An equivalent procedure as applied in the determination of the top background is also considered to determine the uncertainties on the Drell-Yan estimation by the *Transfer Factor* Method. The breakdown of the detector systematic and modelling uncertainties on the Drell-Yan background estimation in the SR, DYCR and on the $\mathrm{TF_{DY}}$ are detailed in Table A.3. The modelling uncertainties are calculated by comparing the nominal MadGraph+Pythia8

sample with two alternative MC samples POWHEG+PYTHIA8 and SHERPA2.2. The dominant uncertainties come from the nominal MC statistics, $E_{\mathrm{T}}^{\mathrm{miss}}$ measurements and also the MC modelling.

The measured $\mathrm{TF}_{\mathrm{DY}}$ has an assigned statistical uncertainty of 14.6% originating from the nominal MC sample statistics while the total uncertainty is 37.6%. The final value of the measured transfer factor is

$$\mathrm{TF}_{\mathrm{DY}} = 0.040 \pm 0.015.$$

The estimated Drell-Yan background in the SR using the *Transfer Factor* method is

$$N_{\mathrm{DY}}^{\mathrm{SR}} = 63 \ \pm 2\,(\mathrm{stat.}) \ \pm 24\,(\mathrm{syst.}).$$

5.4.3. Fake Background

The events from the W+jets production enter the SR when the W boson decays leptonically to a lepton and a neutrino and an associated hadronic jet is misidentified as the other lepton in the $e\mu$ pair. Multi-jet events can also be mistaken for WW events if two hadronic jets are misidentified as an electron and a muon and mismeasurements or fluctuations due to pileup cause a sufficiently large $E_{\mathrm{T}}^{\mathrm{miss}}$. The rate at which hadronic jets are misidentified as leptons may not be accurately described in the MC samples. Therefore, the fake (W+jets and multi-jet) background contribution is determined from data using the data-driven *Matrix Method*.

In this analysis, the two leptons in the data events can either be both real leptons produced by W boson's leptonic decay or at least one of them can originate from a jet misidentified as a lepton, referred to as the fake lepton. The fake lepton may also come from semileptonic decays of heavy-flavour jets or leptons produced by photon conversions. The *Matrix Method* identifies data events with one or two fake leptons to estimate the W+jets and multi-jet contributions respectively.

For simplified demonstration, the applicability of the *Matrix Method* is discussed for a hypothetical sample consisting of events with only a single lepton in the final state. The method is thereafter generalised to estimate the fake background in this analysis where the data events comprise two final state leptons.

Matrix Method for 1-lepton Scenario

The *Matrix Method* compares events satisfying two sets of lepton selection requirements - *loose* (L) and *tight* (T) where the *tight* is a subset of the *loose* selection. For both selections, the number of events in the sample can be represented as the sum of events containg a real and a

fake lepton as shown in the following set of equations.

$$N^{\mathrm{L}} = N^{\mathrm{L}}_{\mathrm{fake}} + N^{\mathrm{L}}_{\mathrm{real}} \tag{5.4}$$
$$N^{\mathrm{T}} = N^{\mathrm{T}}_{\mathrm{fake}} + N^{\mathrm{T}}_{\mathrm{real}} \tag{5.5}$$
$$= \epsilon_{\mathrm{fake}} N^{\mathrm{L}}_{\mathrm{fake}} + \epsilon_{\mathrm{real}} N^{\mathrm{L}}_{\mathrm{real}} \tag{5.6}$$

where $\epsilon_{\mathrm{fake}} = \frac{N^{\mathrm{T}}_{\mathrm{fake}}}{N^{\mathrm{L}}_{\mathrm{fake}}}$ and $\epsilon_{\mathrm{real}} = \frac{N^{\mathrm{T}}_{\mathrm{real}}}{N^{\mathrm{L}}_{\mathrm{real}}}$ are the efficiencies of the fake and real leptons selected with *loose* selection to satisfy also the *tight* requirement.

With known real (ϵ_{real}) and fake (ϵ_{fake}) lepton efficiencies, Equations 5.4 and 5.6 can be solved to obtain the the number of fake events in the *tight* sample as

$$N^{\mathrm{T}}_{\mathrm{fake}} = \frac{\epsilon_{\mathrm{fake}}}{\epsilon_{\mathrm{real}} - \epsilon_{\mathrm{fake}}} (\epsilon_{\mathrm{real}} N^{\mathrm{L}} - N^{\mathrm{T}}) \tag{5.7}$$

Since the efficiencies are dependent on the lepton kinematic properties in an event, the fake contribution is not calculated according to the Equation 5.7 instead each event is assigned with a weight to be either fake or real like. The total fake contribution is therefore measured as the sum of the event weights using the following equation.

$$N^{\mathrm{T}}_{\mathrm{fake}} = \sum_i w_i = \sum_i \frac{\epsilon_{i,\mathrm{fake}}}{\epsilon_{i,\mathrm{real}} - \epsilon_{i,\mathrm{fake}}} (\epsilon_{i,\mathrm{real}} - \delta_i) \tag{5.8}$$

where $\delta_i = 1$ if the lepton passes the *tight* selection and 0 otherwise. Therefore, if the lepton efficiencies are known, the total number of fake events with *tight* selection can easily be calculated from Equation 5.8.

Matrix Method for 2-lepton Scenario

The *Matrix Method* approach as discussed above can now be used to estimate the fake contribution in the selected data events in this analysis. Two alternative sets of lepton selection requirements are likewise used - *loose* (L) and *tight* (T) where the *tight* is a subset of the *loose* requirement. The *tight* selection for electrons and muons are same as used for the SR event selection. The *loose* electrons are selected by relaxing the electron identification requirement and also by removing the isolation requirement from the *tight* sample. The *loose* muons, on the other hand, only require the isolation criterion to be removed from the *tight* muon selection. Both the leptons must fulfil the basic kinematic requirements and one of them must satisfy the assigned single lepton trigger criteria of the analysis before entering the *loose* and *tight* criteria.

The selection sets used in the *Matrix Method* fake background estimation are defined as follows.

- **Electrons:**

  - *loose* (L): Medium identification
  - *tight* (T): Tight identification + Gradient isolation
- **Muons:**
  - *loose* (L): Medium identification
  - *tight* (T): Medium identification + Gradient isolation

The total number of events include events with exactly two *loose* leptons (N^{LL}), one *loose* lepton and one *tight* lepton (N^{LT} and N^{TL}) and two *tight* leptons (N^{TT}). The first and second indices correspond to the qualities of the leading and subleading leptons respectively. On the other hand, the selected *loose* or *tight* lepton might be real or fake. Therefore, the total events can also be visualised into four categogies - (real,real), (real,fake), (fake,real) and (fake,fake). These two sets of categories can be connected to each other in terms of efficiecies of real (ϵ_{real}) and fake (ϵ_{fake}) leptons selected with *loose* criteria to satisfy the *tight* selection. These dependencies are described in the following set of equations.

$$N^{\mathrm{LL}} = N^{\mathrm{LL}}_{\mathrm{fake,fake}} + N^{\mathrm{LL}}_{\mathrm{real,fake}} + N^{\mathrm{LL}}_{\mathrm{fake,real}} + N^{\mathrm{LL}}_{\mathrm{real,real}} \tag{5.9}$$

$$N^{\mathrm{LT}} = \epsilon_{\mathrm{fake}} N^{\mathrm{LL}}_{\mathrm{fake,fake}} + \epsilon_{\mathrm{fake}} N^{\mathrm{LL}}_{\mathrm{real,fake}} + \epsilon_{\mathrm{real}} N^{\mathrm{LL}}_{\mathrm{fake,real}} + \epsilon_{\mathrm{real}} N^{\mathrm{LL}}_{\mathrm{real,real}} \tag{5.10}$$

$$N^{\mathrm{TL}} = \epsilon_{\mathrm{fake}} N^{\mathrm{LL}}_{\mathrm{fake,fake}} + \epsilon_{\mathrm{real}} N^{\mathrm{LL}}_{\mathrm{real,fake}} + \epsilon_{\mathrm{fake}} N^{\mathrm{LL}}_{\mathrm{fake,real}} + \epsilon_{\mathrm{real}} N^{\mathrm{LL}}_{\mathrm{real,real}} \tag{5.11}$$

$$N^{\mathrm{TT}} = \epsilon^2_{\mathrm{fake}} N^{\mathrm{LL}}_{\mathrm{fake,fake}} + \epsilon_{\mathrm{real}}\epsilon_{\mathrm{fake}} N^{\mathrm{LL}}_{\mathrm{real,fake}} + \epsilon_{\mathrm{fake}}\epsilon_{\mathrm{real}} N^{\mathrm{LL}}_{\mathrm{fake,real}} + \epsilon^2_{\mathrm{real}} N^{\mathrm{LL}}_{\mathrm{real,real}}. \tag{5.12}$$

The number of events with two *tight* leptons selected by the analysis SR selection, consist of contributions from multi-jet events with two fake leptons, W+jets events with one fake and one real lepton and finally events with two real leptons including the WW signal events as described in Equation 5.12. If the number of events with *loose* and *tight* leptons as well as the efficiencies ϵ_{real} and ϵ_{fake} are known, the number of events with one real and one fake lepton ($N^{\mathrm{LL}}_{\mathrm{real,fake}}$+$N^{\mathrm{LL}}_{\mathrm{fake,real}}$) and two fake leptons ($N^{\mathrm{LL}}_{\mathrm{fake,fake}}$) for the *loose* selection criteria can be obtained by solving the above system of equations. The number of W+jets and multi-jet events in the signal region, which are selected using the two *tight* leptons, can then be extracted using the following relations.

$$N_{W+\mathrm{jets}} = \epsilon_{\mathrm{real}}\epsilon_{\mathrm{fake}} N^{\mathrm{LL}}_{\mathrm{real,fake}} + \epsilon_{\mathrm{fake}}\epsilon_{\mathrm{real}} N^{\mathrm{LL}}_{\mathrm{fake,real}} \tag{5.13}$$

$$N_{\mathrm{multi-jet}} = \epsilon^2_{\mathrm{fake}} N^{\mathrm{LL}}_{\mathrm{fake,fake}} \tag{5.14}$$

The W+jets and multi-jet contributions are calculated by the sum of the weights assigned per-event to be either (real,fake)+(fake,real) or (fake,fake) like. The efficiencies ϵ_{real} and ϵ_{fake} are determined separately as discussed in the following sections. For the WW signal signature considered in this analysis, the leading and subleading leptons are of different flavour. Therefore, in both Equation 5.13 and 5.14, the product and squared terms of the efficiencies would be the

product of electron and muon efficiencies (real or fake) depending on the lepton to be either leading or subleading.

Measurement of Real Lepton Efficiency

The efficiencies of the real electrons and muons are measured on the WW signal MC sample. The SHERPA (corresponds to SHERPA2.1) MC sample is selected for this purpose instead of the nominal POWHEG (corresponds to POWHEG+PYTHIA8) WW MC. This is preferred as the SHERPA sample offers $\sim$5-6 times larger MC statistics. The MC sample is corrected by applying appropriate Scale Factors (SFs) to account for the data to MC efficiency difference for both *loose* and *tight* lepton selection criteria. The SFs are extracted by taking the ratio of data and MC efficiencies calculated on $Z \to ee/\mu\mu$ events as discussed in Section 4.2 and 4.3. Since the MC has been corrected to reflect the data efficiencies, this procedure is equivalent to a data-driven efficiency measurement.

The quality of the leptons passing the *loose* or *tight* selections depend on the trigger decision imposed on the lepton candidates which in turn profoundly influence the efficiencies. Therefore, the lepton efficiencies are classified into two categories. The 'trig' efficiencies refer to when the lepton (electron or muon) passes the corresponding single lepton trigger requirement (Table 5.2). The 'notrig' efficiencies are calculated for leptons not being able to satisfy the trigger criterion. For both scenarios, the other lepton in the $e\mu$ pair must fulfil the trigger condition as per the nominal analysis trigger selection.

The kinematic bias that might impact the efficiencies is also tested for the leptons being either leading or subleading lepton in one of the two ($e\mu$ or μe) channels.

Figure 5.8 shows the efficiencies of electron and muon, split according to trigger decisions as well as $e\mu$ or μe channels. As expected, there are no conclusive differences seen between leading and subleading lepton efficiencies within statistical fluctuations. Hence, the final efficiencies as input to the *Matrix Method* fake background determination are considered to be the sum of both $e\mu$+μe channels to enhance the statistical power. The $e\mu$+μe real lepton 'trig' and 'notrig' efficiencies are independently treated as inputs given the significant differences seen between the two categories.

Uncertainties on Real Lepton Efficiency

The sources of uncertainties assigned to the real lepton efficiency measurement are listed below.

- **MC Modelling**
  The real lepton efficiencies are measured on the alternative SHERPA WW MC rather the nominal POWHEG. The difference in the measured efficiencies from the two samples is examined. The statistical uncertainty associated with the efficiencies measured from the nominal POWHEG MC are also studied. Figure 5.9 shows the real lepton efficiencies

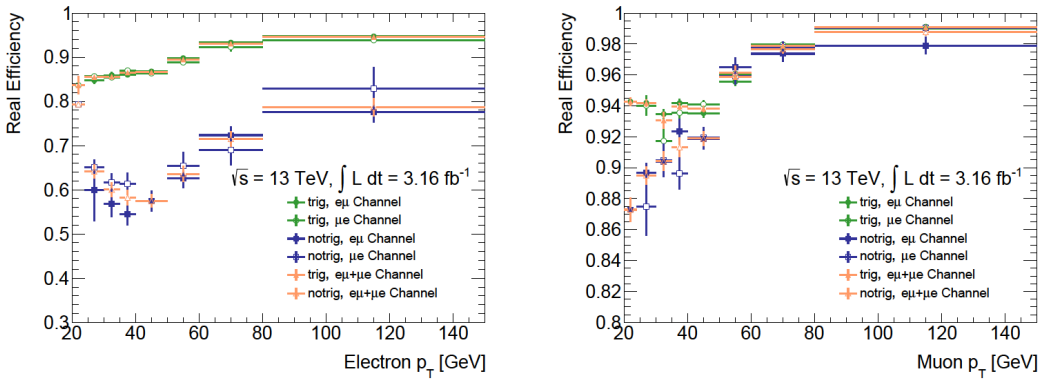


Figure 5.8.: Efficiencies of real electrons (left) and muons (right) as a function of their transverse momentum p_{T}. They are split into $e\mu$ and μe final states as well as 'trig' and 'notrig' categories. The final inputs are $e\mu + \mu e$ channel efficiencies classified based on the trigger decisions. Only the statistical uncertainty originating from the SHERPA WW MC sample statistics is included.

measured in both the SHERPA and POWHEG MCs. The largest of the two contributions is added as the MC modelling uncertainty to the total systematic variation of the efficiency measurement.

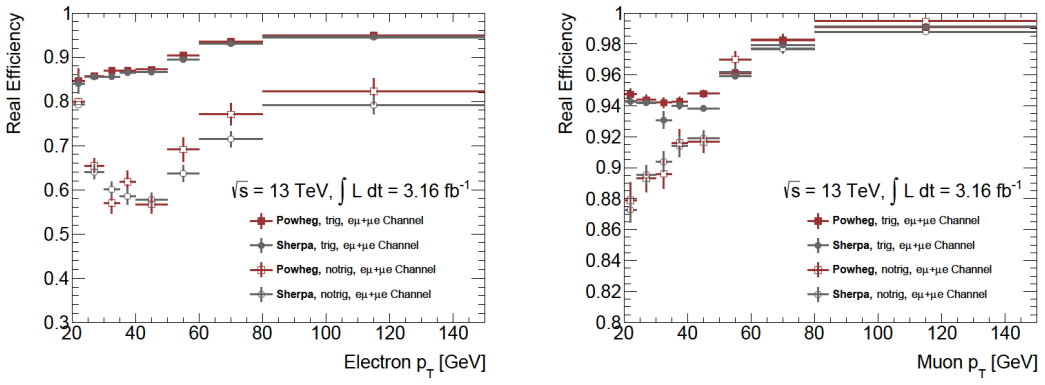


Figure 5.9.: Efficiencies of real electrons (left) and muons (right) as a function of their transverse momentum p_{T}. The efficiencies measured in SHERPA (gray) and POWHEG (red) MC samples are compared for both 'trig' and 'notrig' categories in $e\mu$+μe channel. Only the statistical uncertainty is included.

- **Propagation of Scale Factor Uncertainties**
  The WW MC sample is corrected to data using SFs. The total uncertainty associated with the SF measurements influence the real lepton efficiencies. The propagation of uncertainties related to different SFs affecting the real electron and muon efficiencies are as follows.
  - Electrons: Both identification and isolation SFs.
  - Muons: Only isolation SF.

For electrons, both identification and isolation SFs are calculated by imposing requirements on track variables, calorimetric shower shapes and also measured from highly correlated samples. Hence, identification and isolation SF uncertainties are treated as correlated and added linearly. This linear sum is denoted as the Electron SF uncertainty.

- **Statistical Uncertainty**
  The statistical uncertainty of the efficiency calculation comes from the SHERPA MC sample statistics.

Figure 5.10 and 5.11 present the individual uncertainty sources contributing to the real electron and muon efficiency measurements for 'trig' and 'notrig' categories. The MC modelling, Electron SF and statistical uncertainties are then added in quadrature to measure the total uncertainty on the real electron efficiency. For muon efficiency, only isolation SF contributes and therefore all uncertainty components are added in quadrature to get the final total uncertainty on the real muon efficiency measurement.

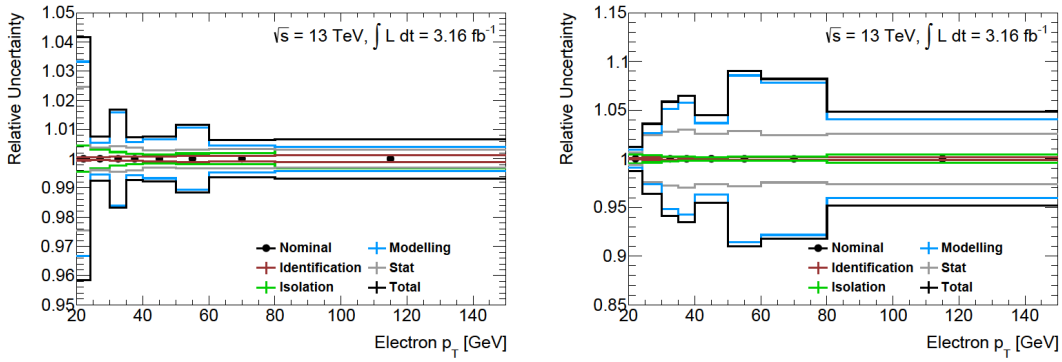


Figure 5.10.: Individual uncertainty sources and the total uncertainty contributing to the real electron 'trig' (left) and 'notrig' (right) efficiencies as a function of electron transverse momentum p_{T}.

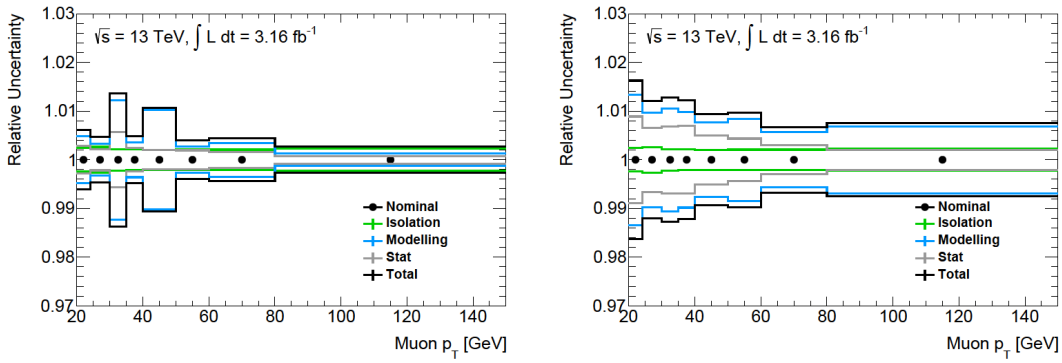


Figure 5.11.: Individual uncertainty sources and the total uncertainty contributing to the real muon 'trig' (left) and 'notrig' (right) efficiencies as a function of muon transverse momentum p_{T}.

Measurement of Fake Lepton Efficiency

The fake lepton efficiencies are measured in the data sample by selecting events that contain exactly one charged lepton and at least one associated hadronic jet. The lepton always satisfies the single-lepton trigger criterion to fulfil the WW analysis selection. The selected data sample with exactly one lepton is chosen to avoid contaminations from real di/multi-lepton, such as WW, other diboson and top events.

The fake lepton efficiencies are calculated [98] in a control region with low lepton transverse mass $\mathrm{m_T(lep)}$ $(= \sqrt{2p_{\mathrm{T}}^{\ell} \cdot p_{\mathrm{T}}^{\mathrm{miss}} \cdot (1 - \cos\Delta\phi)})$ and low $E_{\mathrm{T}}^{\mathrm{miss}}$, enriched with fake lepton contribution. The real lepton contributions from W/Z and top events entering the control region are further subtracted using MC predictions.

The control region uses the following selections:

- $\mathrm{m_T(lep)} < 20$ GeV
- $(\mathrm{m_T(lep)} + E_{\mathrm{T}}^{\mathrm{miss}}) < 60$ GeV

Uncertainties on Fake Lepton Efficiency

The sources of systematic uncertainties contributing to the fake lepton efficiency measurement are described in the following.

- **Alternative Control Region**
  An alternative control region using $E_{\mathrm{T}}^{\mathrm{miss}} < 20$ GeV selection is investigated to study the impact of real lepton contribution (mainly from the Drell-Yan process) on the fake efficiency measurement.

- **Jet Variation**
  The fake lepton efficiencies are highly affected by the presence of the number of hadronic jets in the data events. The jet multiplicity requirement in the events modifies the real lepton contribution primarily from top processes and non-isolated leptons which in turn alters the measured fake efficiency. The fake efficiencies are studied on data events allowing exactly one jet, zero b-jet, one b-jet, one jet - zero b-jet and two jets as variations. The largest deviation coming from these variations in jet composition is added as a systematic effect on the measured fake lepton efficiency.

- **MC Cross Section**
  The real lepton contribution from W/Z and top processes are subtracted using the MC predictions. The cross section of these MC samples is varied (increased or decreased) by 30% [98] to study the possible biases that might be introduced by the MC subtraction. The deviation in the measured fake efficiency due to this MC cross section variation is added as the final source of uncertainty.

The individual systematic variations on the fake lepton efficiencies are shown in Figure 5.12 and 5.13. The uncertainties due to jet variation and alternative control region are symmetrised in both up/down systematics on top of the measured value. The total uncertainty is finally calculated from the quadrature sum of all uncertainty components discussed above.

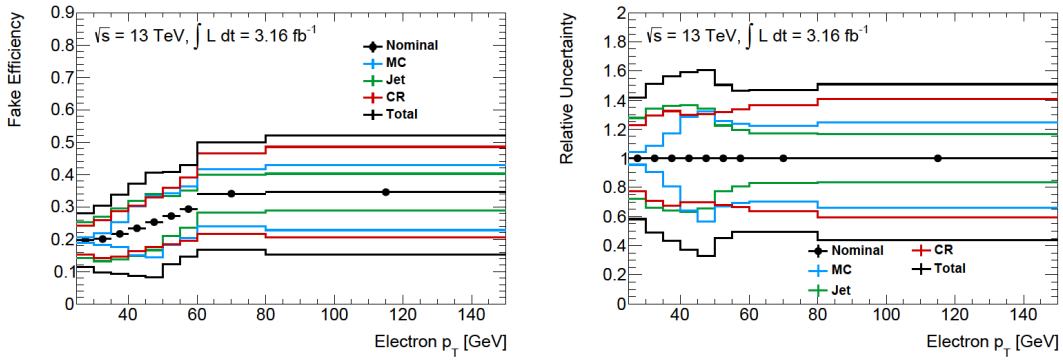


Figure 5.12.: Fake electron efficiencies (left) for nominal and all systematic variations as a function of electron transverse momentum p_{T}. The percentage uncertainty contributions from individual sources and the total uncertainty associated with the fake electron efficiency measurement are shown in the right plot.

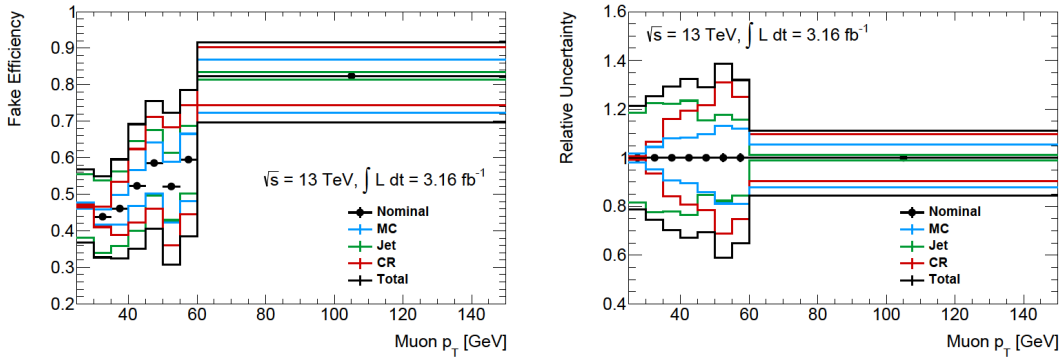


Figure 5.13.: Fake muon efficiencies (left) for nominal and all systematic variations as a function of muon transverse momentum p_{T}. The percentage uncertainty contributions from individual sources and the total uncertainty associated with the fake muon efficiency measurement are shown in the right plot.

Closure Test on W+jets MC

The validity of the *Matrix Method* is tested on the W+jets MC by performing a closure test. Only the W+jets process is studied for this test as it comprises more than 90% (Table 5.11) of the total fake background entering the selected data sample in this analysis. The input fake lepton efficiencies are measured on the W+jets MC samples using the same *loose* and *tight* selection definitions. The real lepton efficiencies are measured from the WW signal MC (Sherpa). Finally, the *Matrix Method* is applied on the W+jets MC similarly as it is applied to

the data. The MC samples used in the efficiency measurements or in the estimation itself are not being corrected to data using Scale Factors as the method is carried out on the MC itself.

Matrix Method Closure Test	W+jets *Matrix Method* prediction		W+jets MC prediction
	p_{T} dependence	η-p_{T} dependence	
$e\mu$ Channel	23.7 ± 4.0	25.2 ± 5.1	39.0 ± 7.9
μe Channel	24.4 ± 3.8	28.9 ± 5.8	19.8 ± 5.8
$e\mu + \mu e$ Channel	48.1 ± 5.5	54.1 ± 7.7	58.7 ± 9.8

Table 5.10.: SR events yields as a result of the closure test performed on the W+jets MC. The W+jets *Matrix Method* estimations are compared to the POWHEG MC prediction for the $e\mu$, μe and combined $e\mu + \mu e$ channels. Only the MC statistical uncertainty is included.

Table 5.10 contains the results of the performed MC closure test where the estimated W+jets background using the *Matrix Method* is compared to the MC prediction of this process in the SR. As the efficiencies could have a large dependence on both η and p_{T} of the leptons, closure tests have been done by taking into account both p_{T} and η-p_{T} dependent efficiencies.

As the result of the closure test, both the *Matrix Method* and MC predictions seem to be compatible with each other given their statistical fluctuations. The statistical uncertainty on the MC prediction comes from the events selected with the *tight* requirement. On the other hand, the *Matrix Method* deals with *loose* events and applies event weights depending on whether the events pass or fail the *tight* selection. This is reflected in the associated statistical uncertainties (Table 5.10) where the uncertainties are smaller in the *Matrix Method* estimation than the MC prediction. The existing deviation between the two predictions mentioned above is considered as the 'non-closure' uncertainty which is assigned to the final data-driven *Matrix Method* estimation for the fake background.

W+jets and Multi-jet Estimation

The *Matrix Method* data-driven (DD) estimations of W+jets and multi-jet events entering the WW SR are listed in Table 5.11. The SR W+jets prediction from POWHEG MC sample is included in the table for comparison.

$e\mu + \mu e$ Channel	W+jets + multi-jet DD	multi-jet DD	W+jets DD	W+jets MC
SR	77.7 ± 19.0	6.7 ± 3.5	71.0 ± 20.8	48.3 ± 9.6

Table 5.11.: Events yields of data-driven (DD) W+jets and multi-jet estimation using *Matrix Method* and W+jets MC prediction in the SR. Only the statistical uncertainty from data or MC sample statistics is included.

The uncertainties associated with the real and fake electron (Figure 5.10 and 5.12) and muon (Figure 5.11 and 5.13) efficiencies are treated as individual source of systematic variations to the

final fake (W+jets and multi-jet) estimation. The up or down uncertainties of each efficiency are varied at a once while the other efficiency inputs are kept fixed to their nominal value. The deviations in the fake estimation are treated as an individual up/down systematic variation. An additional 'non-closure' uncertainty is added as the result of the closure test. Finally, all uncorrelated systematic variations are added in quadrature to obtain the total systematic uncertainties on the fake estimations as summarised in Table 5.12. The statistical uncertainty comes from the data sample statistics.

Source of uncertainty (in %)	$e\mu + \mu e$ Channel 100 %		$e\mu$ Channel 69.6 %		μe Channel 30.4 %	
	Up Var	Down Var	Up Var	Down Var	Up Var	Down Var
Real electron efficiency	+7.2	−7.7	+5.1	−5.6	+11.9	-12.7
Fake electron efficiency	+38.9	−23.7	+12.2	−5.3	+100.0	−65.8
Real muon efficiency	+21.7	−22.9	+13.8	−14.5	+39.9	−42.2
Fake muon efficiency	+16.9	−16.9	+44.8	−30.9	−46.9	+14.9
Total efficiency syst.	48.2	37.8	48.7	35.0	118.0	94.9
Non-closure	22.0		64.6		18.8	
Syst.	53.0	43.7	80.9	73.5	119.5	96.7
Stat.	24.5		24.0		59.0	

Table 5.12.: All sources of systematic uncertainties on the fake background estimation using the data-driven *Matrix Method*. The total up and down systematic uncertainties are the quadrature sum of all four efficiency contributions and the 'non-closure' uncertainty listed above. The 'non-closure' uncertainty has been assigned as the difference in *Matrix Method* and MC predictions from the closure test performed on W+jets MC as described in Table 5.10.

The total number of fake events estimated by the *Matrix Method* in the SR is

$$N_{\text{fakes}}^{\text{SR}} = 78 \pm 19\,(\text{stat.}) \pm {}^{41}_{34}\,(\text{syst.}).$$

Same Sign Validation Region

The validity of the data-driven fake background estimation using the *Matrix Method* is cross-checked in a phase space independent to the SR selection. This validation region is decided by selecting two isolated different flavour leptons of the same charge. The same charge requirement suppresses the WW signal contribution in this region. Requiring no additional third lepton and $m_{e\mu} > 10$ GeV selections bring this region close to the SR definition and also decrease other diboson processes. Furthermore, jet veto and $p_{\text{T}}^{\text{miss}} > 20$ GeV cuts are applied to reduce the top and Drell-Yan events. The resulting region enriched with W+jets events provides a reasonably beneficial cross-check of the *Matrix Method* estimation. Other diboson backgrounds (WZ, ZZ and $V + \gamma$) also contribute significantly in this region.

Figure 5.14 shows the data/MC comparisons of the kinematic distributions related to the di-lepton quantities in this Same Sign Validation Region (SSVR). The SSVR provides a compatible data/MC agreement within the statistical uncertainty and justifies the reliability of the *Matrix Method* approach.

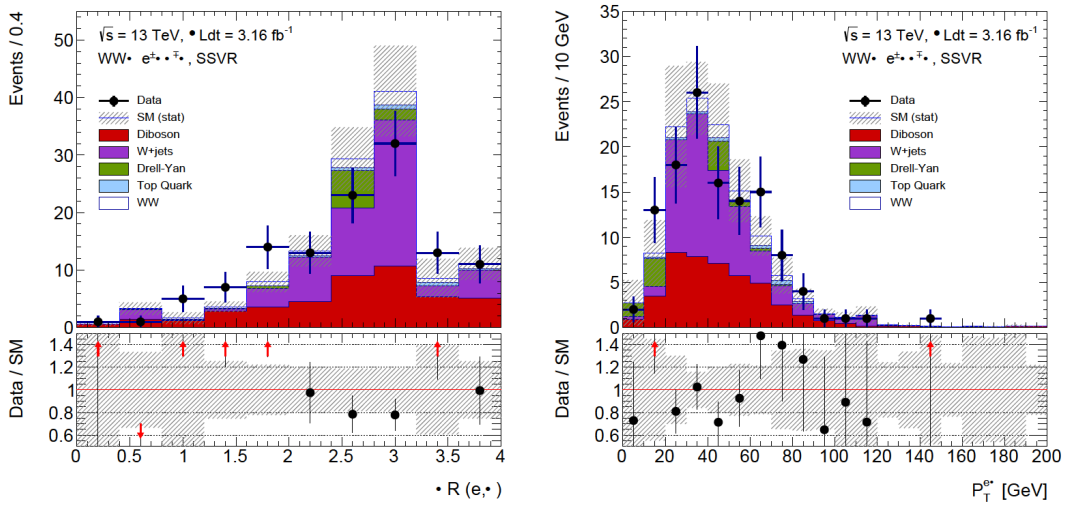


Figure 5.14.: Data/MC comparisons of di-lepton distributions - ΔR between the two leptons (left) and transverse momentum of the $e\mu$ pair (right) in the SSVR. The W+jets *Matrix Method* prediction is included. The uncertainty covers only the statistical component coming from the data/MC sample statistics.

5.4.4. Other Diboson Background

The final background component to enter the WW selection results from the other diboson processes, such as WZ, ZZ and $V+\gamma$ ($V = W, Z$). The contributions of these processes are reduced significantly by requiring no additional leptons to be present in the SR selected events. A subdominant contribution (2% of the total MC prediction in the SR) from these processes persists due to the additional lepton(s) falling outside the detector acceptance or being non-identified by the lepton selection algorithms as discussed in Sections 4.2 and 4.3. These backgrounds are estimated using the corresponding MC predictions. Table 5.13 lists the nominal MC predictions of these processes in the SR.

$e\mu + \mu e$ Channel	WZ Powheg+Pythia8	ZZ Powheg+Pythia8	$WZ + ZZ$	$V+\gamma$ Sherpa 2.1
SR	17.9 ± 0.3	1.1 ± 0.1	18.9 ± 0.3	18.3 ± 1.8

Table 5.13.: MC predictions of WZ, ZZ and $V+\gamma$ processes in the SR. Only the MC statistical uncertainty is included.

Validation Region

The MC predictions for WZ and ZZ processes are validated in a phase space independent to the SR, denoted as the Diboson Validation Region (DibosonVR). This DibosonVR is constructed by allowing events containing one e and one μ as the SR selection. However, no charge requirement is imposed on the leading and subleading leptons. To enhance the WZ and ZZ contributions, an additional third lepton with $p_{\mathrm{T}} > 10$ GeV is allowed to enter the selected sample by dropping the $N^{\ell}_{\mathrm{add}} = 0$ requirement. The third lepton is then matched with the leading or subleading lepton to construct a same flavour lepton pair. To ensure this same flavour lepton pair to originate from a Z boson candidate in the WZ or ZZ process, the invariant mass of this pair is required to be within 15 GeV of the Z boson mass. This selection rejects a part of $t\bar{t}$ contribution in this validation region.

The data/MC comparisons for the di-lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ related distributions in the Diboson VR are shown in Figure 5.15. This region is dominated by the WZ, ZZ, Drell-Yan and top processes. It demonstrates a good agreement of MC with data to provide confidence in the MC predictions for the WZ and ZZ processes.

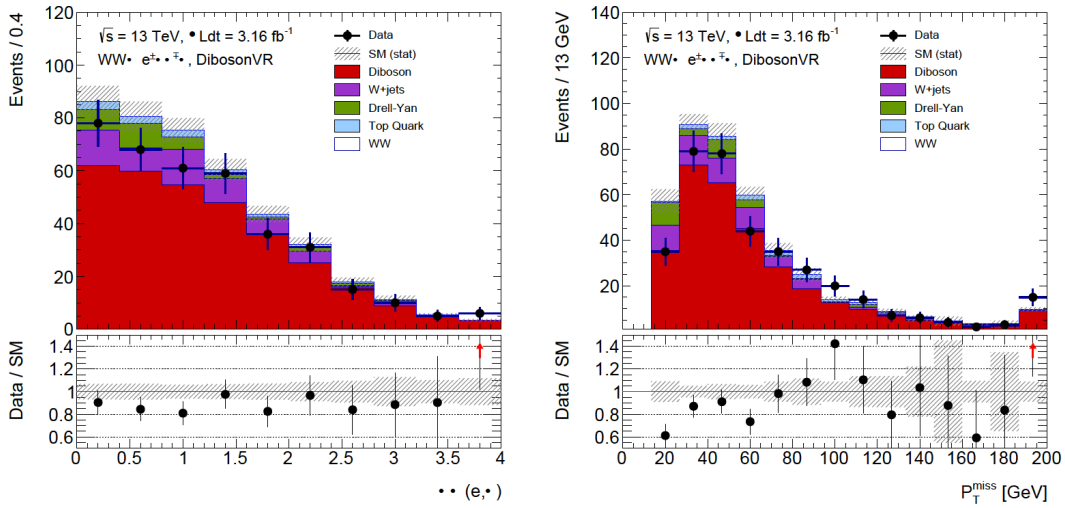


Figure 5.15.: Data/MC comparisons of the kinematic distributions - $\Delta\eta$ between the two leptons in the $e\mu$ pair (left) and $p_{\mathrm{T}}^{\mathrm{miss}}$ (right) in the DibosonVR. The uncertainty covers only the statistical component coming from the data/MC sample statistics.

The other diboson backgrounds in the SR from the MC-based predictions are assigned 17.2% uncertainty where the dominant source of uncertainties come from jet calibrations and MC modelling. The modelling uncertainty of the WZ and ZZ processes are determined by comparing the events yields predicted by the nominal Powheg+Pythia8 with Sherpa2.1 generator. For the $V + \gamma$ process, no other alternative generators were available to compare and hence the same modelling uncertainty assigned to the WZ and ZZ processes is added to

the $V + \gamma$ process as well. The breakdown of the total uncertainties associated with the other diboson background prediction is listed in Table A.4.

The MC predicted other diboson (WZ, ZZ and $V + \gamma$) events entering the SR is

$$N^{\mathrm{SR}}_{\mathrm{other\,diboson}} = 38 \ \pm 2\,(\mathrm{stat.}) \ \pm 6\,(\mathrm{syst.}).$$

Chapter 6.

Cross Section Measurement

This analysis studies the WW production process in the fully leptonic $e\mu$ decay channel. The detector-level event selection to select the observed data and the estimations of signal and background events in the chosen SR phase space have been discussed in the previous Chapter 5. The cross section of the WW process is measured in the particle-level fiducial phase space as described in Section 6.1. The fiducial region eliminates the dependencies on the limited geometrical detector acceptance. The event selection in the fiducial region is preferred to be as close as possible to the detector-level SR selection so that the extrapolation stands out to be minimal between the two phase spaces.

The $pp \to WW$ cross section in the particle-level fiducial phase space is measured using Equation 6.1.

$$\sigma^{\text{fiducial}}_{WW\to e\mu} = \frac{N_{\text{obs}} - N_{\text{bkg}}}{C_{WW} \cdot L} \tag{6.1}$$

where $L = 3.16\,\text{fb}^{-1}$ is the total integrated luminosity of the used dataset. N_{obs} and N_{bkg} are the number of observed data and estimated background events that fulfil the SR selection. The number of WW signal events is given by the difference in $(N_{\text{obs}} - N_{\text{bkg}})$. The C_{WW} factor converts these SR selected events into particle-level expectations. This factor is measured in the MC simulation and accounts for the detector efficiencies in the event reconstruction as well as the response resolution of the detector.

The measured fiducial cross section is then extrapolated to the total phase space. The $pp \to WW$ total cross section is calculated as

$$\sigma^{\text{total}}_{WW} = \frac{\sigma^{\text{fiducial}}_{WW\to e\mu}}{A_{WW} \cdot 2 \cdot \mathcal{B}r^2} = \frac{N_{\text{obs}} - N_{\text{bkg}}}{(A_{WW} \times C_{WW}) \cdot L \cdot 2 \cdot \mathcal{B}r^2} \tag{6.2}$$

where the A_{WW} acceptance factor corrects for the geometric and kinematic restrictions imposed due to the fiducial region selection. The branching ratio $\mathcal{B}r = 0.1083$ [99] is the probability of a W boson to decay into a lepton ($W \to \ell\nu$, $\ell = e, \mu$). Therefore, the $\mathcal{B}r^2$ term

includes both the final state leptons in a WW event candidate. A factor of 2 is considered since the events entering total cross section formula include both the $e\mu$ and μe channels.

Section 6.2 details the calculations of the acceptance factors C_{WW} and A_{WW} in order to proceed to the cross section measurement. The fiducial and total cross sections are measured by the simultaneous fit performed across the signal and the two control regions as discussed in Secion 6.3. Section 6.4 finally summarises the measured cross section results and their comparisons with the theoretical predictions.

6.1. Fiducial Phase Space Definition

The fiducial phase space employs similar event selection cuts as applied to the reconstructed events in the SR. However, it utilises the truth-level MC generated events before passing through the detector simulation. The truth MC events further satisfy the selection requirements in the fiducial region.

The truth electrons and muons are needed to stem from one of the W bosons produced in the hard scattering interaction. The jets in the truth events are formed using the same *anti-k_t* algorithm [62] with the input distance parameter $R = 0.4$. All stable particles including muons, are taken as input to build the truth jet. The jets constructed from the signal leptons are removed from the list of the associated truth jets in an event if they lie within $\Delta R = 0.3$ of a truth lepton. The $E_{\mathrm{T}}^{\mathrm{miss}}$ is estimated by considering the four-vector momentum sum of the neutrinos produced in the W boson's decay.

Table 6.1 lists the truth-level event selection in the fiducial region.

Object requirements for truth-level event selection	
Lepton selection	
Lepton pair selection	$e\mu$ pair, opposite charge
Lepton transverse momentum $p_{\mathrm{T}}^{e,\mu}$	> 25 GeV
Lepton pseudorapidity $\lvert\eta_{e,\mu}\rvert$	< 2.5
Number of additional leptons N_{add}^{ℓ}	$= 0$
Di-lepton invariant mass $m_{e\mu}$	> 10 GeV
Additional object selection	
Number of jets N_{jets}	$= 0$
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 20 GeV
$E_{\mathrm{T,rel}}^{\mathrm{miss}}$	> 15 GeV

Table 6.1.: $WW \to e\mu$ truth-level event selection in the fiducial phase space.

In comparison with the SR selection (Table 5.3), the b-jet veto requirement is dropped. The effect of the b-jet veto cut on the signal events is less than 1% as demonstrated in Table 5.6 and

therefore is not considered in the fiducial region selection. Hence, any source of uncertainties related to the b-tagging procedure is treated to be negligible in the fiducial region.

The WW signal events and the relative compositions of the three WW subprocesses satisfying the fiducial region selection cuts are shown in Table 6.2. This can infer a direct comparison to the SR reconstructed event yields presented in Table 5.6.

$e\mu + \mu e$ Channel $L = 3.16$ fb^{-1}	WW Signal	$q\bar{q} \to WW$		$gg \to WW$		$gg \to H \to WW$	
		$N_{\text{events}}^{\text{truth}}$	Frac.	$N_{\text{events}}^{\text{truth}}$	Frac.	$N_{\text{events}}^{\text{truth}}$	Frac.
Generated truth $e\mu$ events	9429 ± 18	8332 ± 18	0.88	466 ± 3	0.05	631 ± 2	0.07
Lepton selection	3449 ± 11	3144 ± 11	0.91	204 ± 2	0.06	100 ± 1	0.03
Jet veto: $N_{\text{jets}} = 0$	2027 ± 9	1885 ± 8	0.93	108 ± 2	0.05	34 ± 0	0.02
$E_{\text{T}}^{\text{miss}} > 20$ GeV	1605 ± 8	1475 ± 7	0.92	99 ± 2	0.06	32 ± 0	0.02
$E_{\text{T,rel}}^{\text{miss}} > 15$ GeV	1511 ± 7	1385 ± 7	0.92	95 ± 2	0.06	31 ± 0	0.02

Table 6.2.: Total WW signal events and the relative compositions of the events coming from the three signal subprocesses passing each fiducial region selection requirement. The predictions are derived from the nominal signal MC samples scaled by the total luminosity $L = 3.16$ fb^{-1} of the dataset used in this analysis. Only the MC statistical uncertainty is included.

6.2. Acceptance Factors

The following section discusses the calculations of the acceptance factors C_{WW} and A_{WW} in the fiducial and total cross section measurements.

C_{WW} Calculation

The C_{WW} factor occurring in the fiducial cross section definition (Equation 6.1) is measured in the WW signal MC samples using Equation 6.3. It is defined as the ratio of reconstructed signal events passing the full SR selection to the truth-level MC generated signal events fulfilling the fiducial region selection.

$$C_{WW} = \frac{N_{WW\to(e/\mu/\tau)\nu(e/\mu/\tau)\nu\to(e/\mu)\nu(\mu/e)\nu}^{\text{reco, SR}}}{N_{WW\to(e/\mu)\nu(\mu/e)\nu}^{\text{truth, fid}}} \tag{6.3}$$

The numerator of Equation 6.3 selects all reconstructed events containing one electron and one muon in the final state by incorporating all possible leptonic decay modes of the W bosons ($W \to \ell\nu, \ell = e, \mu, \tau$). It implies that in the selected reconstructed events with an $e\mu$ final state, at least one of the leptons may originate from an intermediate τ lepton's decay ($WW \to (e/\mu/\tau)\nu\tau\nu_\tau$). The denominator, on the other hand, only accepts events containing

an $e\mu$ pair where both the leptons are produced from the prompt or direct decay of the two W bosons and excludes the $W \to \tau\nu_\tau$ decay mode. The relative contribution from the $W \to \tau\nu_\tau$ process in the total $WW \to e\mu$ signal events is 8% as estimated in the MC simulation. Therefore, the C_{WW} factor already applies a subtraction of the $W \to \tau\nu_\tau$ contribution based on its relative branching fraction. As a consequence of this, the fiducial cross section measurement in this analysis only considers prompt electron and muon. The effectiveness of this approach relies on the accurate modelling and kinematic properties of the $W \to \tau\nu_\tau$ process. It allows a fair comparison with the previous WW cross section measurement [100] at the centre-of-mass energy of $\sqrt{s} = 8$ TeV using the 2012 ATLAS dataset that used the same strategy.

Uncertainties on C_{WW} Calculation

Different sources of detector and theoretical uncertainties impacting the C_{WW} calculation are listed below.

- **Detector Uncertainties**
  The detector systematic uncertainties influence the reconstructed events entering the C_{WW} calculation. Hence, the effects of these uncertainties on the C_{WW} factor are considered as the sources of detector uncertainties.
- **Parton Shower (PS) Uncertainty**
  The parton shower uncertainty is estimated by comparing the nominal POWHEG+PYTHIA8 MC with an alternative MC sample. The alternative sample is generated with POWHEG matrix element calculation but HERWIG++ parton shower modelling. The difference in the measured C_{WW} factor from both these samples is treated as the parton shower uncertainty.
- **PDF Uncertainty**
  The nominal C_{WW} factor is estimated from the POWHEG+PYTHIA8 sample using the central value of the CT10 PDF set. A variation in C_{WW} is taken by assigning the uncertainty on the CT10 PDF set. The CT10 uncertainty is estimated by varying the input eigenvectors. Two additional variations of the C_{WW} factor are calculated from the generated POWHEG+PYTHIA8 sample but using the central values of the alternative MSTW08 and NNPDF3.0 PDF sets instead of CT10. The largest deviation of these three variations from the nominal C_{WW} is considered as the final PDF uncertainty. The PDF uncertainty calculation mentioned above follows the recommended procedure described in Reference [101].
- **QCD Scale Uncertainties**
  The uncertainties from the higher-order perturbative corrections in QCD are estimated by varying the renormalisation (μ_R) and factorisation (μ_F) scales independently by factors of 0.5 and 2 within the imposed constraint of $0.5 \leq \mu_F/\mu_R \leq 2$. The impact of these variations on the C_{WW} factor is assigned as the QCD scale uncertainty.

The parton shower, PDF and scale uncertainties on the C_{WW} factor are estimated on the dominant $q\bar{q} \to WW$ subprocess with 92% contribution in the total WW signal events selected in both SR (Table 5.6) and fiducial region (Table 6.2) phase spaces. The same uncertainties are further applied to the other two signal subprocesses.

The C_{WW} acceptance factors are calculated individually for the three signal subprocesses from their respective nominal MC samples (Table 5.1). The final value is then derived by combining the C_{WW} factors of these subprocesses based on their contributions in the total WW signal. The measured C_{WW} acceptance factor for the WW signal with its associated statistical and systematic uncertainties is given as

$$C_{WW} = 0.596 \ \pm 0.004\,(\text{stat.}) \ \pm 0.040\,(\text{syst.}).$$

The breakdown of the uncertainties associated with the C_{WW} calculation are summarised in Table 6.3. The dominant uncertainties come from the jet and parton shower variations.

Source of uncertainty	$q\bar{q} \to WW$	$gg \to H \to WW$	$gg \to WW$	Total WW
Stat.	0.6%	1.5%	1.9%	0.6%
Jets	5.5%	9.3%	9.3%	5.9%
Leptons	1.0%	1.3%	1.0%	1.0%
$E_{\mathrm{T}}^{\mathrm{miss}}$	0.2%	0.2%	0.4%	0.2%
Pileup	0.4%	1.3%	0.8%	0.5%
Total Detector	5.6%	9.5%	9.4%	6.0%
PDF		0.2%		
QCD scale		0.2%		
PS		3.1%		
Syst.	6.4%	9.7%	10.0%	6.6%

Table 6.3.: Summary of statistical and systematic uncertainties on the C_{WW} calculation. The jet uncertainty also includes the b-tagging SF variations that affects the reconstructed events satisfying the SR selection.

A_{WW} Calculation

The A_{WW} acceptance factor extrapolates the fiducial cross section into the total phase space. The total phase space contains no kinematic restrictions imposed on it. On the other hand, the fiducial region is defined following the detector-level kinematic and geometric acceptance as well as analysis event selection imposed on the reconstructed events. Therefore, the data and background events involved in the total cross section (Equation 6.2) require corrections from these imposed acceptance restrictions on the fiducial region. This is achieved through the A_{WW} acceptance factor, calculated using Equation 6.4. The A_{WW} factor is determined as the

fraction of all generated $WW \to e\mu$ truth events satisfying the fiducial region selection.

$$A_{WW} = \frac{N^{\text{truth, fid}}_{WW\to(e/\mu)\nu(\mu/e)\nu}}{N^{\text{all truth}}_{WW\to(e/\mu)\nu(\mu/e)\nu}} \tag{6.4}$$

Uncertainties on A_{WW} Calculation

The theoretical uncertainties associated with the A_{WW} calculation are listed below.

- **Parton Shower Uncertainty**
  The parton shower uncertainty is estimated by comparing the nominal POWHEG+PYTHIA8 and the POWHEG+HERWIG++ signal MC samples.

- **PDF Uncertainty**
  The A_{WW} PDF uncertainty is calculated using the same methodology as applied to the C_{WW} factor. The largest deviation to the nominal A_{WW} value coming from the CT10 uncertainty assignment or the central values of any of the two alternatives - MSTW08 and NNPDF3.0 PDF sets is added as the final PDF uncertainty.

- **Jet Veto Uncertainty**
  In a fixed-order cross section measurement, the uncertainty from higher-order corrections in the perturbative series is assessed by the assigned scale uncertainty. The scale uncertainty in such case is calculated by varying the renormalisation (μ_R) and factorisation (μ_F) scales. However, for jet-binned cross section measurement, such as this analysis where the cross section is measured in 0-jet bin using the jet veto requirement, the approach as mentioned earlier may underestimate the scale uncertainty as explained in Reference [102]. This uncertainty underestimation may occur due to accidental cancellations of higher-order perturbative corrections to the total cross section and perturbative corrections that have a logarithmic sensitivity to the jet-bin boundary caused due to the jet veto selection. Therefore, the scale uncertainty is evaluated separately - only for the jet veto selection and other selections together in the fiducial region.

  The jet veto scale uncertainty is calculated by implementing the *Stewart-Tackmann* approach [102]. This approach is based on the assumption that the theoretical scale uncertainties are uncorrelated between different jet multiplicities. The required set of inputs to this method are described below. The cross section in the 0-jet bin σ_0 can be expressed as

  $$\sigma_0 = \sigma_{\text{incl.}} - \sigma_{\geq 1}$$

  where $\sigma_{\text{incl.}}$ and $\sigma_{\geq 1}$ represent the inclusive (no jet multiplicity requirement) and 0-jet exclusive (number of jets ≥ 1) cross sections. The acceptance in σ_0 is

  $$f_0 = \frac{\sigma_{\text{incl.}} - \sigma_{\geq 1}}{\sigma_{\text{incl.}}}$$

Therefore, the uncertainty $\delta(f_0)^2$ on the acceptance f_0 with simple error propagation can be calculated using the following formula.

$$\delta(f_0)^2 = \left(\frac{1}{f_0} - 1\right)^2 (\delta^2_{\text{incl.}} + \delta^2_{\geq 1})$$

- **Scale Uncertainty**
  The part of the scale uncertainty associated with the jet veto selection in the analysis is already considered in the jet veto scale uncertainty. Hence, only the residual uncertainty is evaluated using the standard approach. The renormalisation (μ_R) and factorisation (μ_F) scales are varied by factors of 0.5 and 2 within the constraint of $0.5 \leq \mu_F/\mu_R \leq 2$. This is executed by calculating a partial acceptance factor by considering the full fiducial region selection except for the jet veto requirement, applied to the truth-level MC events.

The aforementioned theoretical uncertainties on the A_{WW} factor are calculated independently for $q\bar{q} \to WW$ and $gg \to H \to WW$ subprocesses. The resulting uncertainties for $gg \to H \to WW$ subprocess are further applied to the $gg \to WW$ signal as well. The measured value of the A_{WW} factor for the WW signal is

$$A_{WW} = 0.1643 \ \pm 0.0010\,(\text{stat.}) \ \pm 0.0093\,(\text{syst.}) = 0.164 \ \pm 0.009\,(\text{tot.}).$$

Table 6.4 summarises the breakdown of the statistical and theoretical uncertainties contributing to the A_{WW} calculation. The largest uncertainties come from the parton shower and jet veto scale uncertainties.

Source of Uncertainty	$q\bar{q} \to WW$	$gg \to H \to WW$	$gg \to WW$	Total WW
Stat.	0.5%	1.0%	1.4%	0.7%
PDF	1.0%	1.0%		1.0%
PS	3.0%	4.0%		3.1%
Jet veto	3.1%	17.2%		4.4%
QCD scale	0.4%	1.4%		0.4%
Syst.	4.5%	17.8%		5.7%

Table 6.4.: Summary of the statistical and theoretical uncertainties on the A_{WW} factor.

6.3. Log-Likelihood Minimisation

As a simplified approach, Equation 6.1 and 6.2 can be used to extract the fiducial and total cross sections of the WW production process. The derived acceptance factors C_{WW} and A_{WW} can be plugged into these cross section formulas along with the constant branching ratio term $\mathcal{B}r^2$ where appropriate. The expected signal events are calculated from the difference in the number of observed data and the estimated background events passing the full SR selection. The backgrounds are either determined from the MC predictions for WZ, ZZ and $V + \gamma$ processes or estimated using dedicated methodologies as detailed in Section 5.4. The fake (W+jets and

multi-jet) background is derived from data using the *Matrix Method*. The top and the Drell-Yan processes are normalised in their respective control regions which are then propagated to the SR using the transfer factors. The extraction of the cross sections using this strategy does not consider any correlations between the signal and background predictions across the signal and control regions. Moreover, it cannot correctly account the statistical fluctuations in these predictions caused due to the finite MC sample statistics.

Therefore, a log-likelihood approach is considered that offers a proper treatment of the correlations between the SR and CR predictions and also the correlations between the systematic uncertainties involved. A negative likelihood function is constructed for that purpose and the procedure is discussed in the following.

The number of expected events in the i^{th} analysis region can be expressed as

$$N^i_{\text{exp}} = N^i_{\text{s}} + N^i_{\text{b}}$$

where N^i_{b} is the number of background events predicted by the MC simulations or estimated using data-driven methods. N^i_{s} represents the number of signal events in the i^{th} analysis region. Given the N^i_{s} signal events, the cross section formalas can be rewritten as

$$N^i_{\text{s}}(\sigma^{\text{fiducial}}_{WW\to e\mu}) = \sigma^{\text{fiducial}}_{WW\to e\mu} \cdot L \cdot C_{WW}$$

$$N^i_{\text{s}}(\sigma^{\text{total}}_{WW}) = \sigma^{\text{total}}_{WW} \cdot 2 \cdot \mathcal{B}r^2 \cdot L \cdot A_{WW} \cdot C_{WW}$$

However, the systematic uncertainties impacting the predicted signal and background events, introduce further corrections. Hence, the expected signal and background events can now be written as

$$N^i_{\text{s}}\left(\sigma^{\text{total}}_{WW}, \{x_k\}\right) = \sigma^{\text{total}}_{WW} \cdot 2 \cdot \mathcal{B}r^2 \cdot L \cdot A_{WW} \cdot C_{WW} \cdot \left(1 + \textstyle\sum_{k=1}^{n} x_k S^i_k\right)$$

$$N^i_{\text{b}}\left(\{x_k\}\right) = N^i_{\text{b}}\left(1 + \textstyle\sum_{k=1}^{n} x_k B^i_k\right)$$

where each systematic uncertainty x_k in the i^{th} analysis region is typically assumed to be a *Normal* distribution with zero mean and unit variance, i.e. $x_k \sim N(0,1)$. The parameters S^i_k and B^i_k are the standard deviation amplitudes, representing the x^{th}_k systematic uncertainty in that region. The expressions are only shown for the total cross section but these apply to the fiducial cross section as well.

In the presence of background normalisation factors, a given prediction is multiplied by such factor without being subjected to any constraint. The following formula shows the number of background events if one of the background processes A has an assigned normalisation factor nf^A.

$$N^i_{\text{b}}\left(\{x_k\}, \text{nf}^A\right) = N^i_{\text{other b}}\left(1 + \textstyle\sum_{k=1}^{n} x_k B^i_k\right) + \text{nf}^A \cdot N^i_A \cdot \left(1 + \textstyle\sum_{k=1}^{n} x_k A^i_k\right)$$

In general, the normalisation factor is obtained in a specific control region with sufficiently high purity of that particular background process of interest. In such a scenario, the systematic uncertainties do not affect the background normalisation in all regions rather the extrapolation between the signal and control regions.

Finally, the negative log-likelihood function can be defined as

$$-\ln L\left(\sigma, \{x_k\}, \mathrm{nf}^j\right) = \sum_i -\ln\left(\frac{e^{-\left(N_\mathrm{s}^i + N_\mathrm{b}^i\right)} \times \left(N_\mathrm{s}^i + N_\mathrm{b}^i\right)^{N_\mathrm{obs}^i}}{(N_\mathrm{obs}^i)!}\right) + \sum_{k=1}^{n} \frac{x_k^2}{2}. \tag{6.5}$$

where the index i runs over the various regions used in the analysis. The expression inside the natural logarithm is essentially the *Poisson* probability that the expected number of signal and background events ($N_\mathrm{s}^i(\sigma, \{x_k\}) + N_\mathrm{b}^i(\{x_k\}, \mathrm{nf}^j)$) will produce the number of events observed in data (N_obs^i) in the i^{th} region. Each k^{th} systematic uncertainty from an independent source is treated as a nuisance parameter x_k and the last term takes care of the *Gaussian* constraints imposed on it. The nuisance parameters are considered to be fully correlated across the regions as well as between the signal and background components. The implementation of this likelihood approach is performed using the ROOStat libraries. Finally, a simultaneous fit across the SR and CRs is performed using the minimisation of the log-likelihood function. The minimisation and the error calculations are performed with the help of the Minuit package [103].

The likelihood approach as discussed above is finally implemented in the context of this analysis. In this analysis, index i runs over three analysis regions - a single SR and two control regions (TopCR and DYCR). The normalisation factors $\mathrm{NF_{top}}$ and $\mathrm{NF_{DY}}$ are assigned to the top and Drell-Yan processes that normalise these backgrounds in their respective CRs.

Results of the Log-Likelihood Minimisation

The final estimates of the signal and background events are obtained as the result of the simultaneous log-likelihood minimisation fit performed across the SR and two CRs. The likelihood function with nuisance parameters automatically takes all the statistical and systematic uncertainties into account and propagates them to the final uncertainty.

The obtained normalisation factors for the top and Drell-Yan processes are reported in Table 6.5. Hence, the MC prediction of the top or Drell-Yan background is required to be scaled by the respective normalisation factor. The signal and other diboson backgrounds are also re-evaluated with a proper correlation treatment of the associated uncertainties. However, the fake background remains unchanged as it is directly derived from data.

Normalisation Factor	Fitted Value
NF_{DY}	1.034 ± 0.030
NF_{top}	0.875 ± 0.035

Table 6.5.: Normalisation factors for the top and Drell-Yan processes as the outcome of the log-likelihood minimisation fit.

Table 6.6 compiles the number of events observed in data as well as the final estimates of the signal and background events after the full SR event selection as the outcome of the performed fit. Estimated events are shown for the signal region as well as the top and Drell-Yan control regions.

Process	Signal Region	Top Control Region	Drell-Yan Control Region
WW signal	997 ± 69	49 ± 12	75.3 ± 5.4
Drell-Yan	62 ± 23	49 ± 29	1568 ± 45
$t\bar{t}$+Wt	177 ± 33	2057 ± 81	3.5 ± 1.6
W+jets/multi-jet	78 ± 41	70 ± 55	0
Other dibosons	38 ± 12	6.3 ± 3.5	19.2 ± 6.1
Total	1351 ± 37	2232 ± 47	1666 ± 41
Data	1351	2232	1666

Table 6.6.: Observed number of events in data and estimated number of events from signal and background processes in the signal and control regions. Events from signal and background processes are the results of the log-likelihood minimisation fit. Both statistical and systematic uncertainties are included.

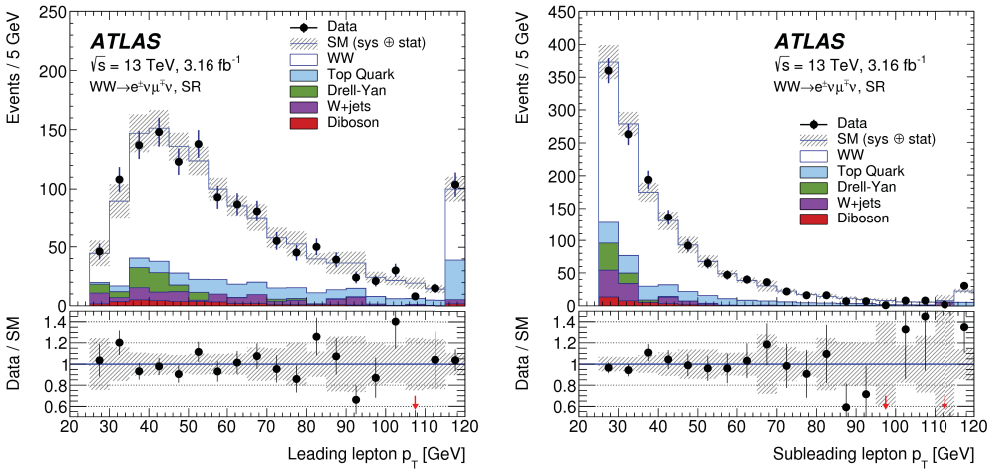
The data/MC comparisons of the kinematic distributions of the lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ related quantities considering the results of the log-likelihood minimisation fit are shown in Figure 6.1 for the SR selection. The corresponding comparisons in the top and Drell-Yan control regions are presented in Figure 6.2. The uncertainty band includes the MC statistical uncertainty as well as full systematic uncertainties associated with the shape and normalisation of the MC samples.

6.4. Results

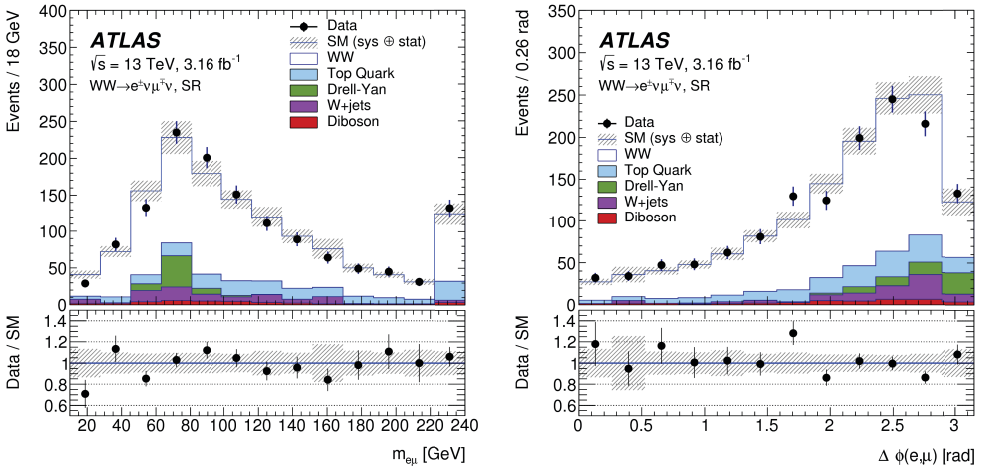
The measured signal cross section in the fiducial phase space as the outcome of the simultaneous log-likelihood minimisation fit is

$$\sigma_{WW\to e\mu}^{\mathrm{fiducial}} = 529 \pm 20\,(\mathrm{stat.}) \pm 50\,(\mathrm{syst.}) \pm 11\,(\mathrm{lumi.})\,\mathrm{fb}.$$

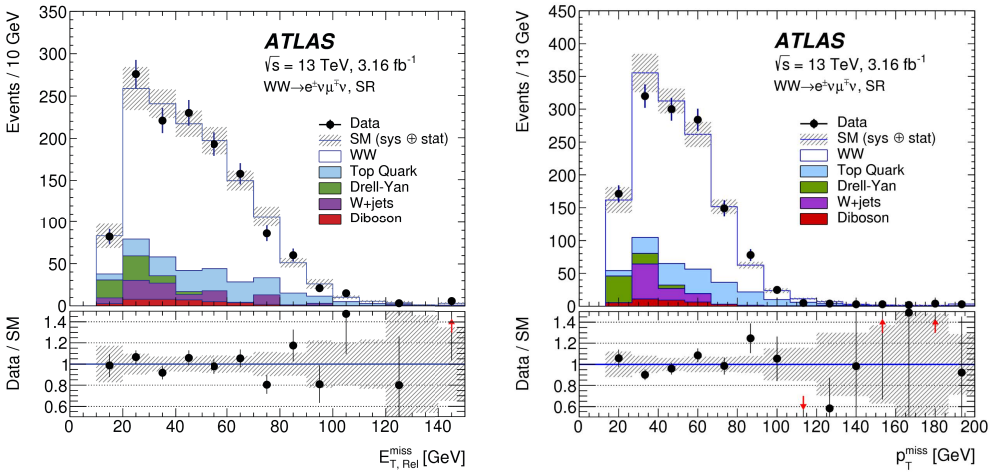
The breakdown of the relative uncertainties associated with the measured fiducial cross section is summarised in Table 6.7. The dominant source of uncertainties originate from the jet selections



(a) Transverse momentum p_{T} of the leading (left) and subleading (right) leptons in the signal region.

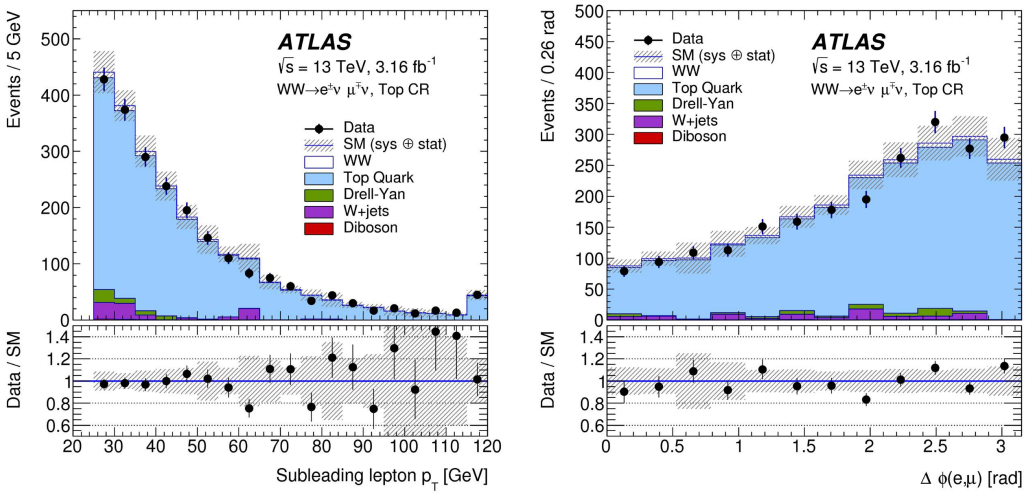


(b) Invariant mass of the $e\mu$ pair (left) and $\Delta\phi$ between the leptons in the $e\mu$ pair (right) in the signal region.

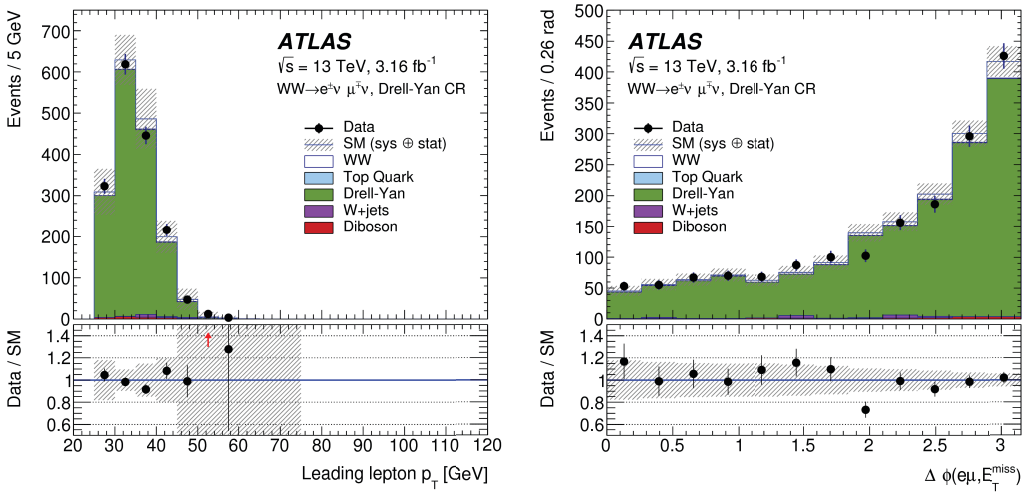


(c) Relative missing transverse momentum $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ (left) and track-based missing transverse momentum $p_{\mathrm{T}}^{\mathrm{miss}}$ (right) in the signal region.

Figure 6.1.: Data/MC comparisons of the kinematic distributions of the final state leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$ related quantities passing the full SR event selection. The signal and background events are the results of the log-likelihood minimisation fit. The uncertainty band includes both the statistical and systematic uncertainties [104].



(a) Transverse momentum p_{T} of the subleading lepton (left) and $\Delta\phi$ between the leptons in the $e\mu$ pair (right) in the top control region.



(b) Transverse momentum p_{T} of the subleading lepton (left) and $\Delta\phi$ between the leptons in the $e\mu$ pair (right) in the Drell-Yan control region.

Figure 6.2.: Data/MC comparisons of the kinematic distributions of the final state leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$ related quantities in the top and Drell-Yan control regions. The signal and background events are the results of the log-likelihood minimisation fit. The uncertainty band includes both the statistical and systematic uncertainties [104].

Sources of uncertainty	Relative uncertainty for $\sigma^{\text{fid}}_{WW\to e\mu}$
Jet selection and energy scale & resolution	7.3%
b-tagging	1.3%
$E_{\text{T}}^{\text{miss}}$ and $p_{\text{T}}^{\text{miss}}$	1.7%
Electron	1.0%
Muon	0.4%
Pileup	0.9%
Luminosity	2.1%
Top background	2.4%
Drell-Yan background	1.5%
W+jet and multi-jet background	3.8%
Other diboson backgrounds	1.1%
Parton shower	3.1%
PDF	0.2%
QCD scale	0.2%
MC statistics	1.2%
Data statistics	3.7%
Total uncertainty	11%

Table 6.7.: Breakdown of the relative uncertainties on the fiducial cross section measurement as a result of the simultaneous fit performed across the signal and control regions.

and calibrations. The measured fiducial cross section is finally compared with the theoretical cross section prediction, discussed in the following.

Theoretical Predictions

The theoretical prediction of the WW cross section is calculated in the total phase space that combines the individual contributions of $q\bar{q} \to WW$, non-resonant $gg \to WW$ and resonant $gg \to H \to WW$ subprocesses. The relative contributions of the $q\bar{q}$, non-resonant and resonant gg productions to the total WW cross section are 87%, 5% and 8% respectively. The cross sections of these subprocesses are known at different $\mathcal{O}(\alpha_{\text{s}})$ as the best available predictions at the time of the measurement. The $q\bar{q} \to WW$ production cross section is known at $\mathcal{O}(\alpha_{\text{s}}^2)$ (NNLO) [105, 106], the non-resonant $gg \to WW$ subprocess is known at $\mathcal{O}(\alpha_{\text{s}}^3)$ (NLO) [107, 108] whereas the resonant $gg \to H \to WW$ cross section is calculated at $\mathcal{O}(\alpha_{\text{s}}^5)$ (N^{3}LO) [109] taking the $H \to WW$ branching fraction [110] into account. These three cross section predictions are combined and the sum is denoted as nNNLO+H prediction. In the nNNLO+H calculation, the interference between the subprocesses is neglected. In the given $\mathcal{O}(\alpha_{\text{s}})$, the $q\bar{q}$ and gg-initiated processes do not interfere. However, an interference effect exists between the two gg-initiated processes - $gg \to WW$ and $gg \to H \to WW$ but the effect of this interference on the total cross section is negligible.

The scale and PDF uncertainties on the theoretical predictions for the WW subprocesses are estimated using the same procedures as applied for the acceptance factor calculations, discussed in Section 6.2. In the total nNNLO+H prediction, the scale uncertainties of different subprocesses are treated as correlated and added linearly. However, the PDF uncertainties are considered to be uncorrelated and added in quadratures.

To pursue a direct comparison with the measured fiducial cross section, the theoretical predictions are also calculated in the fiducial phase space after applying all fiducial region selections. A correction of 0.972 ± 0.001 is applied to the theoretical parton-level cross section prediction to account for the non-perturbative effects. This factor is measured in the POWHEG+PYTHIA8 signal MC sample where the events selected in the fiducial region are compared with and without these non-perturbative effects implemented in the MC sample. The calculations as presented here, do not include higher order electroweak correction. However, in Reference [111], it is estimated that the electroweak correction at NLO would decrease the WW cross section by 3-4%. For $q\bar{q}$ and non-resonant gg productions, the fiducial cross sections are measured by the programs presented in References [106,107]. In case of $gg \to H \to WW$ process, no such fiducial region calculation was available at $\mathcal{O}(\alpha_s^5)$. The $gg \to H \to WW$ fiducial cross section is therefore calculated by applying the A_{WW} acceptance factor measured in the POWHEG+PYTHIA8 MC sample to the inverted total cross section formula 6.2, given as

$$\sigma^{\text{fiducial}}_{WW\to e\mu} = 2 \cdot \sigma^{\text{total}}_{WW} \cdot A_{WW} \cdot \mathcal{B}r^2$$

The uncertainties on the fiducial cross section of the $gg \to H \to WW$ process include the scale and PDF uncertainties for both σ^{total}_{WW} and A_{WW} whereas the parton shower uncertainty is only estimated for A_{WW}. For $q\bar{q}$ and $gg \to WW$ processes, the predicted A_{WW} factors are calculated by their respective fiducial to the total cross section ratios. The scale and PDF uncertainties are treated to be correlated between the fiducial and total cross sections. The dominant uncertainty on theoretical predictions comes from the PDF uncertainty on the A_{WW} factor.

Table 6.8 summarises the total and fiducial cross section predictions for individual WW subprocesses with their assigned uncertainties. This table also reports the predicted A_{WW} factors for the three signal contributions.

Figure 6.3 compares the nNNLO+H fiducial cross section prediction with the measured value of $\sigma^{\text{fiducial}}_{WW\to e\mu} = 529 \pm 20\,(\text{stat.}) \pm 50\,(\text{syst.}) \pm 11\,(\text{lumi.})$ fb. The nNNLO+H prediction of the total cross section is corrected using two alternative A_{WW} calculations to predict the WW cross section in the fiducial phase space. The fixed-order acceptance value is presented in Table 6.8 while the MC acceptance calculation is described in Section 6.2. The nNNLO+H prediction agrees within uncertainties with the experimental cross section measurement in the fiducial phase space.

$pp \to WW$ subprocess	Order of α_s	σ^{tot}_{WW} [pb]	A_{WW} [%]	$\sigma^{fid}_{WW\to e\mu}$ [fb]
$q\bar{q}$ [105, 106]	$\mathcal{O}(\alpha_s^2)$	111.1 ± 2.8	16.20 ± 0.13	422^{+12}_{-11}
gg [107, 108] (non-resonant)	$\mathcal{O}(\alpha_s^3)$	$6.82^{+0.42}_{-0.55}$	$28.1^{+2.7}_{-2.3}$	44.9 ± 7.2
$gg \to H \to WW$ [109, 110]	$\mathcal{O}(\alpha_s^5)$ tot. / $\mathcal{O}(\alpha_s^3)$ fid.	$10.45^{+0.61}_{-0.79}$	4.5 ± 0.6	11.0 ± 2.1
$q\bar{q}$ + gg (non-resonant) + $gg \to H \to WW$	nNNLO+H	$128.4^{+3.5}_{-3.8}$	$15.87^{+0.17}_{-0.14}$	478 ± 17

Table 6.8.: Theoretical predictions for the total and fiducial cross sections for WW subprocesses and their associated uncertainties calculated up to the given order in α_s together with the respective acceptance factors (A_{WW}).

The experimental value of the total cross section is obtained by extrapolating the measured fiducial cross section using the A_{WW} factor from the nNNLO+H prediction (Table 6.8). The calculated value of $\sigma^{total}_{WW} = 142 \pm 5\,(stat.) \pm 13\,(syst.) \pm 3\,(lumi.)$ pb is in agreement with the nNNLO+H theoretical prediction of $128^{+3.5}_{-3.8}$ pb.

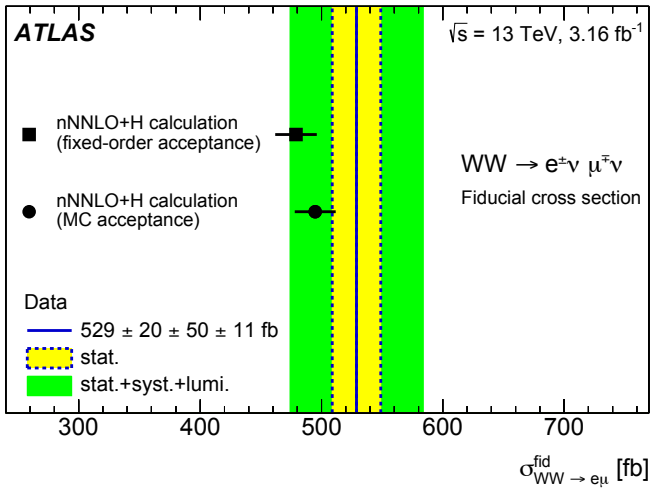


Figure 6.3.: Measured fiducial cross section in comparison with the nNNLO+H prediction in the fiducial phase space with two different acceptance calculations. The vertical bands around the measurement indicate the statistical uncertainty (yellow) and the quadrature sum of statistical, systematic and luminosity uncertainties (green) [104].

Chapter 7.

Sensitivity to Anomalous Couplings

The measured cross section of the WW production process (Section 6.4) within the assigned uncertainty is in good agreement with the SM prediction. In this chapter, the Triple Gauge Boson Couplings (TGCs) - $WW\gamma$ and WWZ deviating from their SM expectations are studied. The anomalous Triple Gauge Boson Couplings (aTGCs) (Section 2.5) represents these deviations of the TGCs from their SM predicted values and are further probed in the WW production process. The non-vanishing aTGC parameters would increase the WW production cross section at higher energies. These would also alter sensitive kinematic distributions of the final state particles, such as p_{T} of the leading lepton towards the higher values. Since the observed data agrees with the SM prediction, the WW production process is further used to constraint physics beyond the Standard Model (BSM) by extracting limits on these aTGC parameters.

The limits are obtained by studying the most sensitive detector-level kinematic distributions of the WW final state particles. A profile-likelihood ratio test statistic [112] is performed to examine the compatibility of the observed data and the predictions with non-zero aTGCs. The 95% Confidence Level (CL) intervals for the five independent aTGC parameters - $\Delta g_1^Z, \Delta k^Z, \lambda^Z, \Delta k^\gamma$ and λ^γ parametrised without any imposed constraint and form-factor (Section 2.5) are determined using the frequentist approach [113]. The confidence interval for a concerned aTGC parameter is calculated when the other four parameters are kept fixed to their SM values.

The data events in this study are selected utilising the same signal region event selection applied to the same dataset as used for the cross section measurement. The WW signal in this case only includes the $q\bar{q} \to WW$ subprocess due to the involvement of the $WW\gamma$ and WWZ couplings (Feynman diagram 2.3a). Therefore, the resonant and non-resonant gg productions estimated using the nominal MC predictions are considered as backgrounds. The $q\bar{q} \to WW$ signal is modelled using the MC@NLO MC sample interfaced with Herwig+Jimmy [114, 115]. A reweighting method [116] implemented in the MC@NLO sample predicts the $q\bar{q} \to WW$ signal events for arbitrary values of the aTGC parameters. This reweighting method utilises an event weight to predict the rate at which the event would be generated in the presence of the anomalous couplings. This event weight is calculated as the ratio of the matrix element squared with (BSM) and without (SM) the aTGC parameters. The background contributions in the

selected data events are estimated with the same methodologies as discussed in Section 5.4. The top and Drell-Yan nominal MC predictions are scaled by the resulting normalisation factors from the log-likelihood minimisation fit performed across the SR and two CRs for the cross section measurement (Table 6.5).

The nominal POWHEG+PYTHIA8 MC sample used to model the $q\bar{q} \to WW$ process for the cross section measurement is generated at NLO and has been normalised to NNLO prediction in the perturbative QCD (Table 5.1) to best describe the expected signal in the selected data. The MC@NLO sample, on the other hand, is generated at NLO in the perturbative QCD calculation. The NLO $\to$ NNLO normalisation applied to the POWHEG+PYTHIA8 MC sample might not be applicable for the MC@NLO sample simulated with non-vanishing aTGC parameters. Therefore, another MC@NLO MC sample is simulated with the SM settings (all aTGCs are set to zero) in which the effect of this NLO $\to$ NNLO normalisation is studied. The contribution originating from this normalisation on the SM-like MC@NLO MC sample is also considered as background.

The kinematic distributions - p_{T} of the leading lepton, invariant mass and p_{T} of the $e\mu$ pair are investigated to extract the limits on the aTGC parameters. The sensitivity towards the aTGC parameters is much more profound in the leading lepton p_{T} distribution in comparison with the other two distributions. The binning used for the leading lepton p_{T} distribution is the same optimised binning that had been used to extract the aTGC limits in the previous WW measurement performed using the 2012 ATLAS dataset at the centre-of-mass collision energy of $\sqrt{s} = 8$ TeV [5].

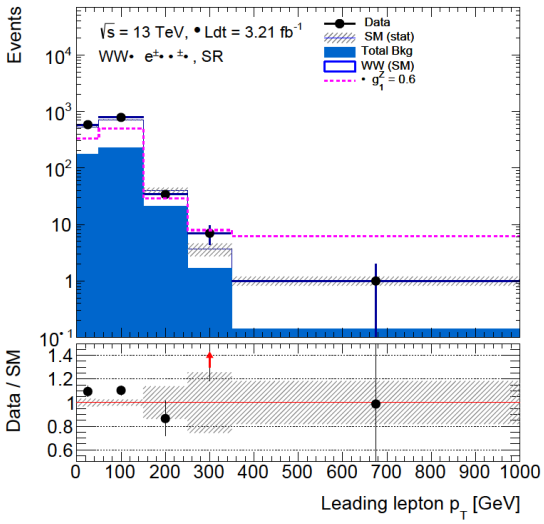


Figure 7.1.: Transverse momentum p_{T} distribution of the leading lepton used to set limits on the aTGC parameters. The observed data, nominal $q\bar{q} \to WW$ signal and background predictions are compared with the MC@NLO MC sample simulated with the aTGC parameter $\Delta g_1^Z = 0.6$. Only the MC statistical uncertainty is included.

Figure 7.1 shows the leading lepton p_{T} distribution presenting the selected data and nominal signal (POWHEG+PYTHIA8) and background predictions using this 'standard binning' (Ta-

ble 7.1). The MC@NLO sample simulated with the aTGC $\Delta g_1^Z = 0.6$ is also overlaid. At the high p_{T} region, MC@NLO prediction with a non-vanishing aTGC underestimates the data. This binning is referred to as the 'standard binning' in the following.

The extracted 95% confidence level intervals for the five independent aTGC parameters based on the leading lepton p_{T} distribution are reported in the following section.

Confidence Intervals

The confidence intervals for the aTGCs based on the leading lepton p_{T} distribution is obtained for the 'standard binning' as mentioned earlier as well as with more number of differential bins, referred to as the 'fine binning' in the following. Table 7.1 lists the two sets of binning of the leading lepton p_{T} distribution in which the aTGC limits are extracted. The 'fine binning' is chosen by looking at the distribution but without performing any further investigations to get the best optimised and sensitive selection. The statistical and systematic uncertainties are calculated differentially in all considered bins of the concerned distribution.

Leading Lepton p_{T} Distribution	
Standard Binning	[0., 50., 150., 250., 350., 1000.] GeV
Fine Binning	[0., 30., 40., 50., 80., 100., 150., 200., 250., 300., 350., 500., 1000.] GeV

Table 7.1.: Binning of the leading lepton p_{T} distribution used in the aTGC limit extraction. The 'standard binning' had been used for the aTGC study in the previous WW measurement with the 2012 ATLAS dataset at the centre-of-mass collision energy of $\sqrt{s} = 8$ TeV [5].

The 95% confidence level intervals for the five independent aTGCs at the cutoff scale $\Lambda_{\mathrm{FF}} = \infty$ are determined for the aforementioned two sets of binning of the leading lepton p_{T} distribution. The limits for a concerned aTGC parameter are calculated while keeping the other parameters fixed to their SM expectations. The observed and Asimov [112] datasets are used to extract the observed and expected limits respectively. The observed limits on the Δg_1^Z are improved by using the 'fine binning' in comparison with the 'standard binning' while for the other parameters, they get slightly worse. The expected limits, on the other hand, get better with more differential bins under consideration. Therefore, an optimal scan over the leading lepton p_{T} distribution to optimise the chosen binning show prospects of higher sensitivity in determining the aTGC parameters. In the following investigations, the results are obtained using the 'standard binning'. The extracted limits based on the other two distributions - invariant mass $m_{e\mu}$ and transverse momentum $P_T^{e\mu}$ of the $e\mu$ pair are presented in Appendix A.3.

The observed and expected aTGC limits are also determined with the combined ATLAS datasets collected in the year of 2012 and 2015. The 2012 ATLAS dataset corresponds to pp collisions at the centre-of-mass energy of $\sqrt{s} = 8$ TeV and an integrated luminosity of $L = 20.3$

Leading lepton p_{T}	Standard Binning		Fine Binning	
$\Lambda_{\mathrm{FF}} = \infty$	Expected	Observed	Expected	Observed
Δg_1^Z	[-0.339, 0.430]	[-0.482, 0.524]	[-0.315, 0.403]	[-0.385, 0.493]
Δk^Z	[-0.047, 0.060]	[-0.057, 0.061]	[-0.043, 0.056]	[-0.059, 0.073]
λ^Z	[-0.038, 0.038]	[-0.044, 0.043]	[-0.035, 0.035]	[-0.047, 0.047]
Δk^γ	[-0.101, 0.130]	[-0.114, 0.119]	[-0.091, 0.119]	[-0.124, 0.150]
λ^γ	[-0.082, 0.082]	[-0.087, 0.084]	[-0.074, 0.075]	[-0.099, 0.099]

Table 7.2.: The 95% CL expected and observed limits on the five independent aTGC parameters with the form factor $\Lambda_{\mathrm{FF}} = \infty$. The limits are extracted for two sets of binning of the leading lepton p_{T} distribution with the 2015 ATLAS dataset used in this analysis. The dataset corresponds to the centre-of-mass collision energy of $\sqrt{s} = 13$ TeV and comprises an integrated luminosity of $L = 3.16$ fb^{-1}.

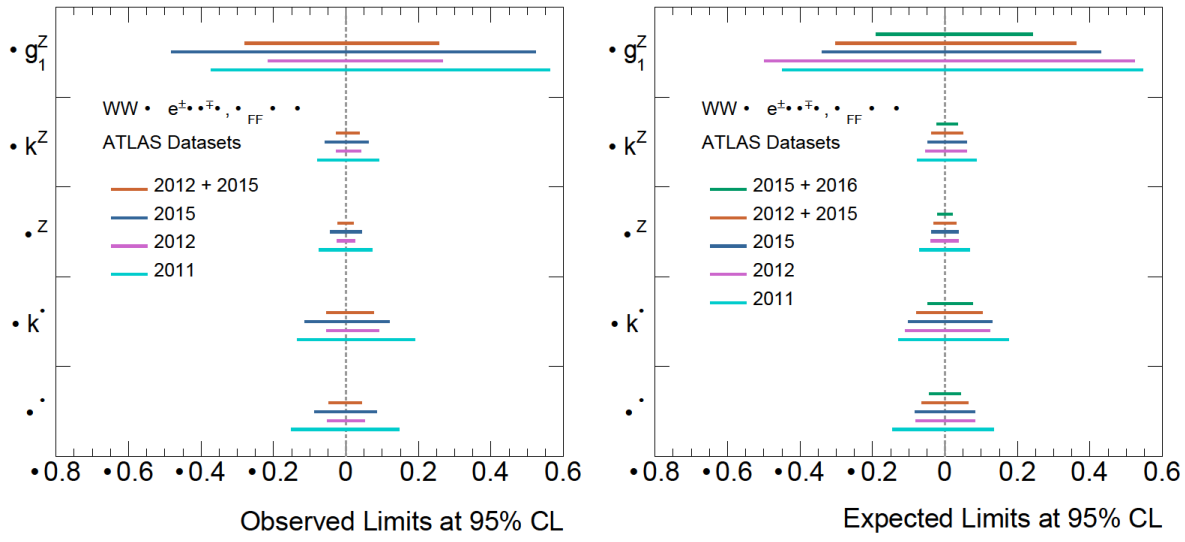
fb^{-1}. The 2015 dataset used in this analysis is collected in the pp collisions at the collision energy of $\sqrt{s} = 13$ TeV and comprises an integrated luminosity of $L = 3.16$ fb^{-1}. The expected limits are also determined for the full luminosity projection of the combined 2015 and 2016 ($\sqrt{s} = 13$ TeV, $L = 33.3$ fb^{-1}) datasets. The results are presented in Table 7.3. The outcomes of these dataset combinations show a significant improvement in the extracted limits in comparison with 2015 dataset alone.

Leading lepton p_{T}	2012 + 2015 Datasets		2015 + 2016 Datasets
$\Lambda_{\mathrm{FF}} = \infty$	Expected	Observed	Expected
Δg_1^Z	[-0.303, 0.362]	[-0.280, 0.257]	[-0.190, 0.243]
Δk^Z	[-0.038, 0.049]	[-0.026, 0.037]	[-0.023, 0.035]
λ^Z	[-0.032, 0.030]	[-0.023, 0.021]	[-0.020, 0.020]
Δk^γ	[-0.079, 0.103]	[-0.053, 0.076]	[-0.048, 0.076]
λ^γ	[-0.065, 0.064]	[-0.047, 0.044]	[-0.044, 0.044]

Table 7.3.: The 95% CL expected and observed limits on the five independent aTGC parameters with the form factor $\Lambda_{\mathrm{FF}} = \infty$. The observed and expected limits are shown for the combined 2012+2015 ATLAS datasets. The expected limits are also presented for the 2015+2016 ATLAS datasets.

Figures 7.2a and 7.2b summarise the observed and expected 95% confidence level intervals extracted for the five aTGC parameters with the form factor $\Lambda_{\mathrm{FF}} = \infty$. The obtained limits using the ATLAS datasets collected at different years of running are compared. The extracted limits with the dataset collected by the ATLAS detector in 2011 at the centre-of-mass collision energy of $\sqrt{s} = 7$ TeV and an integrated luminosity of $L = 4.6$ fb^{-1} are also added for comparison [4].

The aTGC studies presented in this chapter affirm a visible improvement (notably the expected limits with combined 2015 and 2016 datasets) in constraining the aTGC parameters if a large amount of data is taken into consideration. The larger dataset provides more statistical power in order to perform a proper binning optimisation for the sensitivity study of the aTGC parameters as well as removes any statistical fluctuations that might impact the results.



(a) (b)

Figure 7.2.: Observed (left) and expected (right) limits at the 95% CL for the five independent aTGC parameters with $\Lambda_{\mathrm{FF}} = \infty$. The comparisons of the extracted limits using the 2011 [4], 2012 [5], 2015, 2012+2015 and 2015+2016 (only expected limits) ATLAS datasets are shown. Only limits extracted for the WW production process are presented.

Chapter 8.

Summary

The WW production is a challenging process enabling us to test the fundamental gauge structure of the electroweak theory in the Standard Model. The cross section measurements of this process at the ATLAS and CMS experiments at the centre-of-mass energies of 7 and 8 TeV [4–7] consistently observed discrepancies with the theoretical predictions. The differences between the measurements and the predictions are reduced by incorporating higher order theoretical calculations in the perturbative QCD. Therefore, a precise measurement of the WW production process at the collision energy of 13 TeV would confront the best available theoretical prediction to test our understanding.

In this thesis, the measurement of the WW production cross section with the proton-proton collision data collected by the ATLAS detector at the centre-of-mass collision energy of $\sqrt{s} = 13$ TeV has been presented. The dataset corresponds to the total integrated luminosity of $L = 3.16$ fb^{-1}. The WW signal in this analysis contains one electron, one muon and two associated neutrinos in the final state. The dominant background contributions in the WW candidate sample come from the top, Drell-Yan and W+jets and multi-jet processes that exhibit similar final state signature as the WW signal. Other diboson processes, such as WZ, ZZ and $V + \gamma$ enter the WW signal selection as minor backgrounds. The cross section is measured in the fiducial phase space close to the detector acceptance. The measured fiducial cross section of $\sigma^{\text{fiducial}}_{W^+W^-} = 529 \pm 20\,(\text{stat.}) \pm 50\,(\text{syst.}) \pm 11\,(\text{lumi.})$ fb within the assigned uncertainty is compatible with best available SM prediction of 478 ± 17 fb. The systematic uncertainty drives this measurement where the major systematic uncertainty originates from jet energy scale and calibration. The total cross section has also been calculated by extrapolating the measured cross section in the fiducial phase space to the total phase space using the acceptance factors. This WW production cross section measurement has been published in Reference [104].

Moreover, the observed data are further examined for the presence of the anomalous triple gauge boson couplings. The limits are determined for these aTGC parameters. The extracted 95% confidence level intervals of the aTGC parameters are compatible with the SM expectations but are not the most stringent limits assigned so far. However, further investigations with a much larger dataset and a proper treatment of the binning optimisation of the kinematic

distribution of interest demonstrate possible significant improvements of the sensitivity towards the aTGC parameters to constraint physics beyond the Standard Model.

Appendix A.

Auxilary Materials

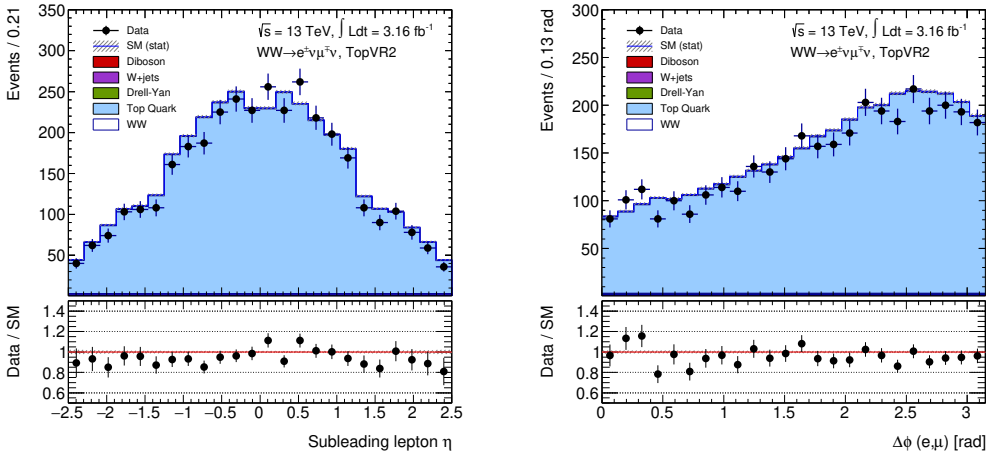
A.1. Additional Top Validation Regions

A further attempt is made to justify the choice of the nominal MC sample used to predict the top events in this analysis. The top control (TopCR) and validation (TopVR) regions (Subsection 5.4.1) select events with a small jet activity which makes these regions close to the SR phase space. The data/MC disagreements in TopCR and TopVR are 11% and 8% respectively. Therefore, the data/MC compatibility is further investigated in two other validation regions (TopVR2 and TopVR3) where the top is the most dominant process with negligible contributions from other processes. Most of the top events mainly originating from $t\bar{t}$ process, contain at least two jets as each top quark produces a jet (mainly b-jet) and a W boson. These regions are constructed using the same lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ selections as in the TopCR and TopVR but with different jet requirements as listed in Table A.1.

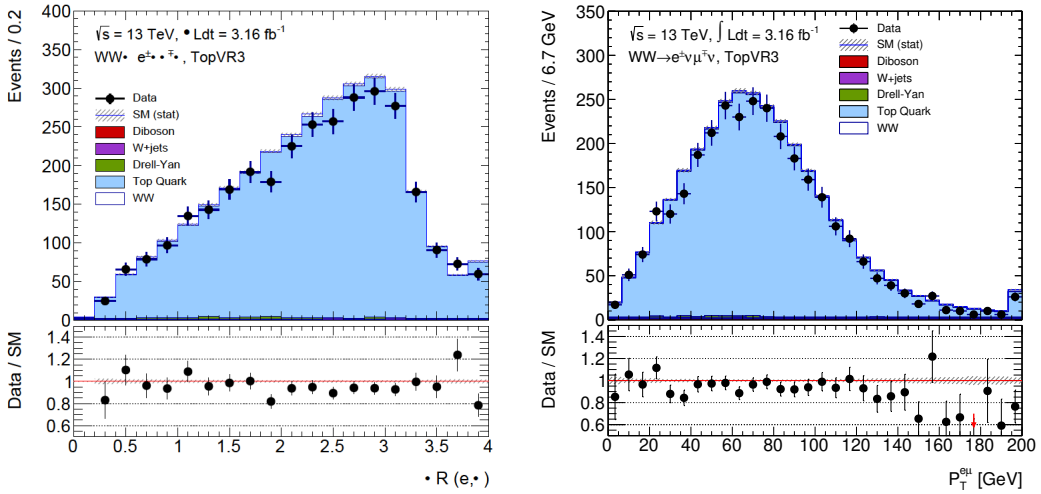
Validation Regions	Jet Selections	Top Purity	Data/MC
TopVR2	$N_{\text{jets}} = 2$, $N_{\text{b-jets}} \geq 1$	0.95	0.93 ± 0.02
TopVR3	$N_{\text{jets}} = 2$, $N_{\text{b-jets}} \geq 2$	0.95	0.97 ± 0.02

Table A.1.: Additional validation regions to check the nominal top MC modelling. The normalisation factor obtained in the TopCR has not been applied to the top MC prediction. Only the statistical uncertainty is included.

The data/MC comparisons of the kinematic distributions of lepton related quantities in both these validation regions are presented in Figure A.1. These regions show an improved data/MC agreement specifically in the validation region TopVR3 constructed with two jets and at least two b-jets requirements. These regions well justify the validity of the nominal top MC modelling.



(a) Pseudorapidity η of the subleading lepton (left) and $\Delta\phi$ between the two leptons in the $e\mu$ pair (right) in TopVR2.



(b) ΔR between the two leptons in the $e\mu$ pair (left) and transverse momentum p_{T} of the $e\mu$ system (right) in the TopVR3.

Figure A.1.: Data/MC comparisons of the kinematic distributions of lepton related quantities in the TopVR2 and TopVR3. The normalisation factor obtained in the TopCR has not been applied to the top MC prediction. Only the statistical uncertainty is included.

A.2. Uncertainties on the Background Estimations

The breakdown of the statistical and systematic uncertainties associated with the background estimations are detailed in the section. Table A.2 and A.3 list the uncertainties on the top and Drell-Yan background estimations using the *Transfer Factor* method. The uncertainties on the other diboson backgrounds estimated from the MC-based predictions are summarised in Table A.4. In these tables, the correlation between different sources of uncertainties are not considered and the quadrature sum of all sources results in the total uncertainty associated with the background estimations.

Top Background	SR	TopCR	TF_{Top}
MC Stat.	±0.9	±0.3	±0.9
Bjet SF	±2.8	±2.7	-5.3/+5.7
Cjet SF	±0.1	±0	±0.1
Light jet SF	±0.6	+0.4/-0.5	-1/+1.1
Electron energy scale	-0.1	±0	±0.1
Electron energy resolution	+0.2/-0.3	±0.2	0/-0.1
Electron reconstruction SF	±0.3	±0.3	±0
Electron identification SF	±0.8	±0.8	±0
Electron isolation SF	±0.3	±0.3	±0
Electron trigger SF	±0.1	±0.1	±0
Jet energy resolution	+2.4	-0.9	+3.3
Jet energy scale NP1	-18.2/+21.9	-13.6/+14.8	-5.3/+6.2
Jet energy scale NP2	-5.9/+6.1	-4.4/+4.5	±1.6
Jet energy scale NP3	-6.8/+7.7	-5.1/+5.2	-1.8/+2.3
Jet vertex tagger	±2.9	±1.8	±1.1
MET TST resolution (1)	±0	-0.2	+0.2
MET TST resolution (2)	±0	-0.1	+0.1
MET TST scale	-0/-0.1	±0.2	+0.1/-0.2
Muon momentum scale	±0	±0	±0
Muon momentum ID resolution	±0	±0	±0
Muon momentum MS resolution	±0.1	-0.1	±0
Muon reconstruction SF	±0.4	±0.4	±0
Muon isolation SF	±0.2	±0.2	±0
Muon trigger SF	±0.1	±0.1	±0
$t\bar{t}$Mod_PS	+2.5	-9.2	+8.8
$t\bar{t}$Mod_ME	+4.1	+11.2	-3.6
$t\bar{t}$Mod_ISR	-5.5	+0.6	-5.1
WtMod_DS/DR	+0.9	+1.3	+0.9
WtMod_PS	-1.8	-4	+2.3
WtMod_ISR	-2.7	+0.3	-2.9
TopXsec_$t\bar{t}$	±3.9	±4.8	±0.8
TopXsec_Wt	±3.5	±2.1	±1.4
Pileup	-0.2/+0.3	-0.5/+0.3	+0.2
Total	±24.6	±22.7	±14.8

Table A.2.: Breakdown of individual sources of uncertainties (expressed in %) on the top background estimation using the *Transfer Factor* method in the SR, TopCR and on the TF_{top}. The up and down uncertainties are averaged to get a single value for each uncertainty source and the total uncertainty is calculated by adding all sources in quadrature. No attempt is made to consider the correlation between the sources at this stage.

Drell-Yan Background	SR	DYCR	TF_{DY}
MC Stat.	± 14.3	± 3	± 14.6
Bjet SF	± 0	± 0	± 0
Cjet SF	± 0.1	± 0.1	± 0
Light jet SF	± 0.5	± 0.2	± 0.2
Electron energy scale	0.2/-1	1.1/-0.4	-1/-0.6
Electron energy resolution	-1.6/-0.2	-0.4/0.3	-1.2/-0.5
Electron reconstruction SF	± 0.3	± 0.3	± 0
Electron identification SF	± 0.9	± 1	± 0.1
Electron isolation SF	± 0.2	± 0.2	± 0
Electron trigger SF	± 0.2	± 0.1	± 0.1
Jet energy resolution	-3.5	-0.7	-2.7
Jet energy scale NP1	-4.5/4.4	-3.3/4.4	-1.3
Jet energy scale NP2	-1.9/3.8	-1.2/1.9	-0.6/1.9
Jet energy scale NP3	-1.5/2.7	-1.1/1.8	-0.4/0.9
Jet vertex tagger	± 4.7	± 0.8	3.9/-3.8
MET TST resolution (1)	-16.3/10.6	0.6/-0.4	-16.8/11.1
MET TST resolution (2)	-14.5	+0.4	-14.8
MET TST scale	-17.4	+0.5	-17.8
Muon momentum scale	± 0	-0.3/0.2	0.3/-0.2
Muon momentum ID resolution	-0.1	-0.2/-0.4	0.1/0.4
Muon momentum MS resolution	-1.6	-0.3/-0.8	-1.3/0.8
Muon reconstruction SF	± 0.4	± 0.4	± 0.1
Muon isolation SF	± 0.2	± 0.2	± 0
Muon trigger SF	± 0.2	± 0.3	± 0
MadGraph vs Powheg	+26.2	+1.8	+9
MadGraph vs Sherpa2.2	+18	-7.6	+18.1
Pileup	3/5.6	-0.8	3.9/5.6
Total	± 37.5	± 9.7	± 37.6

Table A.3.: Breakdown of individual sources of uncertainties (expressed in %) on the Drell-Yan background estimation using the *Transfer Factor* method in the SR, DYCR and on the TF_{DY}. The up and down uncertainties are averaged to get a single value for each uncertainty source and the total uncertainty is calculated by adding all sources in quadrature. No attempt is made to consider the correlation between the sources at this stage.

Other Diboson Background	SR	TopCR	TF_{Top}
MC Stat.	± 4.3	± 11.1	± 7.6
Bjet SF	± 0	± 0.6	± 0
Cjet SF	± 0.1	± 8.9	± 0.1
Light jet SF	± 0.4	+10.6/-10.7	± 0.3
Electron energy scale	+1.2/-0.1	+0.1/-27.4	-1.4
Electron energy resolution	-0.6/+1	-2.3/-27.4	+0.1/-1.2
Electron reconstruction SF	± 0.3	± 0.4	± 0.3
Electron identification SF	± 0.8	± 1	± 1.1
Electron isolation SF	± 0.3	± 0.3	± 0.3
Electron trigger SF	± 0.1	± 0.1	± 0.1
Jet energy resolution	+4.9	-34.1	-1.7
Jet energy scale NP1	-5.4/+13.3	-5.6/+0.8	-4/+2.1
Jet energy scale NP2	-1.8/+2.8	-0.5/-0.8	-0.9/+0.8
Jet energy scale NP3	-2.1/+3.5	-1.3/+3.4	-1/+0.6
MET TST resolution (1)	-2	-5.5	-0.3
MET TST resolution (2)	-1.6	-0.6	+0.8
MET TST scale	-1.1/+1.4	-0.5/+0.1	-0.6/-0.8
Muon momentum scale	-0.2	-0.1	± 0
Muon momentum ID resolution	-0.1/-0.1	-0.2	-1/-0.3
Muon momentum MS resolution	+0.1/-0.2	± 0	+1.3/-0.7
Muon reconstruction SF	± 0.4	± 0.3	± 0.4
Muon isolation SF	± 0.2	± 0.2	± 0.2
Muon trigger SF	± 0.3	± 0.8	+0.6/-0.7
Sherpa vs Powheg	-11.9	+11.1	-11
Pileup	-1/+1.6	-0.1/+1.6	-7.5/+4.6
Total	± 17.2	± 50.9	± 15.4

Table A.4.: Breakdown of individual sources of uncertainties (expressed in %) on the MC-based other diboson background prediction in the SR, TopCR and DYCR. The up and down uncertainties are averaged to get a single value for each uncertainty source and the total uncertainty is calculated by adding all sources in quadrature. No attempt is made to consider the correlation between the sources at this stage.

A.3. Investigations on the Anomalous Couplings

This section includes the extracted 95% confidence level intervals for the aTGC parameters based on the distributions - invariant mass $m_{e\mu}$ and transverse momentum $P_T^{e\mu}$ of the $e\mu$ pair. Table A.5 lists the 'standard binning' of these two distributions which had been optimised for the aTGC limit extraction study in the previous WW measurement with 2012 ATLAS dataset at the centre-of-mass collision energy of $\sqrt{s} = 8$ TeV [5].

	Standard Binning
$m_{e\mu}$	[0., 175., 300., 550., 700., 1000.] GeV
$P_T^{e\mu}$	[0., 125., 225., 300., 1000.] GeV

Table A.5.: The 'standard binning' of the invariant mass $m_{e\mu}$ and transverse momentum $P_T^{e\mu}$ of the $e\mu$ pair used in the aTGC limit extraction. The 'standard binning' refers to the binning used for the aTGC limit extraction in the previous WW measurement with the 2012 ATLAS dataset collected at the centre-of-mass collision energy of $\sqrt{s} = 8$ TeV [5].

The observed and expected 95% confidence level intervals for the five independent aTGC parameters based on these two distributions are reported in Table A.6.

$\Lambda_{\mathrm{FF}} = \infty$	$m_{e\mu}$		$P_T^{e\mu}$	
	Expected	Observed	Expected	Observed
Δg_1^Z	[-0.385, 0.490]	[-0.539, 0.529]	[-0.414, 0.504]	[-0.647, 0.633]
Δk^Z	[-0.062, 0.074]	[-0.071, 0.074]	[-0.054, 0.064]	[-0.066, 0.068]
λ^Z	[-0.048, 0.049]	[-0.053, 0.052]	[-0.042, 0.042]	[-0.049, 0.048]
Δk^γ	[-0.121, 0.147]	[-0.135, 0.144]	[-0.112, 0.134]	[-0.140, 0.144]
λ^γ	[-0.094, 0.095]	[-0.102, 0.099]	[-0.087, 0.088]	[-0.103, 0.100]

Table A.6.: The 95% CL expected and observed limits on the five independent aTGC parameters with the form factor $\Lambda_{\mathrm{FF}} = \infty$. The limits are extracted for two kinematic distributions - invariant mass $m_{e\mu}$ and transverse momentum $P_T^{e\mu}$ of the $e\mu$ pair with the 2015 ATLAS dataset used in this analysis. The dataset corresponds to the centre-of-mass collision energy of $\sqrt{s} = 13$ TeV and comprises an integrated luminosity of $L = 3.16$ fb^{-1}.

Bibliography

[1] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv:1207.7214 [hep-ex].

[2] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv:1207.7235 [hep-ex].

[3] Particle Data Group Collaboration, *Review of Particle Physics*, Chin. Phys. C **40** (2016) 100001.

[4] ATLAS Collaboration, *Measurement of W^+W^- production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector and limits on anomalous WWZ and $WW\gamma$ couplings*, Phys. Rev. D **87** (2013) 112001, arXiv:1210.2979 [hep-ex], [Erratum: Phys. Rev. D **88** (2013) 079906].

[5] ATLAS Collaboration, *Measurement of total and differential W^+W^- production cross sections in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector and limits on anomalous triple-gauge-boson couplings*, J. High Energy Phys. **09** (2016) 029, arXiv:1603.01702 [hep-ex].

[6] CMS Collaboration, *Measurement of the W^+W^- cross section in pp collisions at $\sqrt{s} = 7$ TeV and limits on anomalous $WW\gamma$ and WWZ couplings*, Eur. Phys. J. C **73** (2013) 2610, arXiv:1306.1126 [hep-ex].

[7] CMS Collaboration, *Measurement of the $\mathrm{W^+W^-}$ cross section in pp collisions at $\sqrt{s} = 8$ TeV and limits on anomalous gauge couplings*, Eur. Phys. J. C **76** (2016) 401, arXiv:1507.03268 [hep-ex].

[8] T. Gehrmann et al., *W^+W^- production at hadron colliders in NNLO QCD*, Phys. Rev. Lett. **113** (2014) 212001, arXiv:1408.5243 [hep-ph].

[9] P. Jaiswal and T. Okui, *Explanation of the WW excess at the LHC by jet-veto resummation*, Phys. Rev. D **90** (2014) 073009, arXiv:1407.4537 [hep-ph].

[10] M. Grazzini et al., *Transverse-momentum resummation for vector-boson pair production at NNLL+NNLO*, J. High Energy Phys. **08** (2015) 154, arXiv:1507.02565 [hep-ph].

[11] P. Meade et al., *Transverse momentum resummation effects in W^+W^- measurements*,

Phys. Rev. D **90** (2014) 114006, arXiv:1407.4481 [hep-ph].

[12] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **22** (1961) 579.

[13] S. L. Glashow et al., *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D **2** (1970) 1285.

[14] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13** (1964) 168.

[15] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[16] G. 't Hooft and M. Veltman, *Regularization and renormalization of gauge fields*, Nucl. Phys. B **44** (1972) 189.

[17] D. J. Griffiths, *Introduction to elementary particles; 2nd rev. version*, Physics textbook, Wiley, (2008).

[18] E. Noether, *Invariant Variation Problems*, Gott. Nachr. **1918** (1918) 235, Trans. Theo. Stat. Phys. 1 (1971) 186, arXiv:physics/0503066.

[19] Super-Kamiokande Collaboration, *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562, arXiv:hep-ex/9807003.

[20] M. Gell-Mann, *The interpretation of the new particles as displaced charge multiplets*, Nuovo Cim. **4** (1956) 848.

[21] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321.

[22] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508.

[23] K. G. Wilson, *Confinement of quarks*, Phys. Rev. D **10** (1974) 2445.

[24] M. N. Chernodub, *Background magnetic field stabilizes QCD string against breaking*, arXiv:1001.0570 [hep-ph].

[25] S. Bethke et al., *World Summary of α_s (2015)*, EPJ Web Conf. **120** (2016) 07005.

[26] G. 't Hooft, *Dimensional regularization and the renormalization group*, Nucl. Phys. B **61** (1973) 455.

[27] J. C. Collins et al., *Factorization of Hard Processes in QCD*, Adv. Ser. Direct. High Energy Phys. **5** (1989) 1, arXiv:hep-ph/0409313.

[28] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nucl. Phys. B **126** (1977) 298.

[29] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024.

[30] S. Dulat et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, Phys. Rev. D **93** (2016) 033006.

[31] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, J. High Energy Phys. **07** (2002) 012, arXiv:hep-ph/0201195.

[32] A. D. Martin et al., *Parton distributions for the LHC*, Eur. Phys. J. C **63** (2009) 189, arXiv:0901.0002 [hep-ph].

[33] B. Andersson et al., *Parton Fragmentation and String Dynamics*, Phys. Rept. **97** (1983) 31.

[34] B. R. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, Nucl. Phys. B **238** (1984) 492.

[35] GEANT4 Collaboration, *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250.

[36] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823, arXiv:1005.4568 [physics.ins-det].

[37] R. W. Brown and K. O. Mikaelian, W^+W^- *and* Z^0Z^0 *pair production in* e^+e^-, pp, *and* $\overline{p}p$ *colliding beams*, Phys. Rev. D **19** (1979) 922.

[38] ALEPH, DELPHI, L3, OPAL Collaborations, LEP Electroweak Working Group, *Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP*, Phys. Rept. **532** (2013) 119, arXiv:1302.3415 [hep-ex].

[39] K. Hagiwara et al., *Low energy effects of new interactions in the electroweak boson sector*, Phys. Rev. D **48** (1993) 2182.

[40] H. Aihara et al., *Anomalous gauge boson interactions*, arXiv:hep-ph/9503425.

[41] C. Degrande et al., *Effective Field Theory: A Modern Approach to Anomalous Couplings*, Annals Phys. **335** (2013) 21, arXiv:1205.4231 [hep-ph].

[42] L. Evans and P. Bryant, *LHC Machine*, J. Instrum. **3** (2008) S08001.

[43] ATLAS Collaboration, *ATLAS public results of luminosity measurement in Run 2*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[44] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, J. Instrum. **3** (2008) S08003.

[45] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report, 1*, CERN-LHCC-97-016, ATLAS-TDR-4, (1997), https://cds.cern.ch/record/331063.

[46] ATLAS Collaboration, *Studies of the ATLAS Inner Detector material using* $\sqrt{s}$ *=13 TeV pp collision data*, ATL-PHYS-PUB-2015-050, (2015), http://cds.cern.ch/record/2109010.

[47] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, CERN-LHCC-2010-013, ATLAS-TDR-19, (2010), https://cds.cern.ch/record/1291633.

[48] ATLAS Collaboration, *ATLAS Inner Detector Alignment Performance with February 2015 Cosmic Rays Data*, ATL-PHYS-PUB-2015-009, (2015), https://cds.cern.ch/record/2008724.

[49] ATLAS Collaboration, *ATLAS liquid-argon calorimeter: Technical Design Report*, CERN-LHCC-96-041, ATLAS-TDR-2, (1996), https://cds.cern.ch/record/331061.

[50] ATLAS Collaboration, *ATLAS muon spectrometer: Technical Design Report*, CERN-LHCC-97-022, ATLAS-TDR-10, (1997), https://cds.cern.ch/record/331068.

[51] ATLAS Collaboration, *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*, CERN-LHCC-2003-022, ATLAS-TDR-16, (2003), https://cds.cern.ch/record/616089.

[52] ATLAS Collaboration, *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System*, CERN-LHCC-2013-018, ATLAS-TDR-023, (2013), https://cds.cern.ch/record/1602235.

[53] ATLAS Collaboration, *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002, (2007), https://cds.cern.ch/record/1020106.

[54] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. A **262** (1987) 444.

[55] ATLAS Collaboration, *Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2*, Eur. Phys. J. C **77** (2017) 673, arXiv:1704.07983 [hep-ex].

[56] ATLAS Collaboration, *Performance of primary vertex reconstruction in proton-proton collisions at* $\sqrt{s} = 7$ *TeV in the ATLAS experiment*, ATLAS-CONF-2010-069, (2010), https://cds.cern.ch/record/1281344.

[57] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, ATLAS-CONF-2016-024, (2016), http://cds.cern.ch/record/2157687.

[58] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, (2008), https://cds.cern.ch/record/1099735.

[59] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, ATLAS-CONF-2012-047, (2012),

https://cds.cern.ch/record/1449796.

[60] ATLAS Collaboration, *Electron Efficiency Measurements in 2015 Data*, https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/EGAM-2015-006/index.html.

[61] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton-proton collision data at* $\sqrt{s}$ *=13 TeV*, Eur. Phys. J. C **76** (2016) 292, arXiv:1603.05598 [hep-ex].

[62] M. Cacciari, G. P. Salam, G. Soyez, *The anti-*k_t *jet clustering algorithm*, J. High Energy Phys. **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[63] ATLAS Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at* $\sqrt{s} = 13$ *TeV*, ATL-PHYS-PUB-2015-015, (2015), https://cds.cern.ch/record/2037613.

[64] ATLAS Collaboration, *Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS*, ATLAS-CONF-2015-037, (2015), https://cds.cern.ch/record/2044941.

[65] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018, (2014), https://cds.cern.ch/record/1700870.

[66] ATLAS Collaboration, *Expected performance of the ATLAS b-tagging algorithms in Run-2*, ATL-PHYS-PUB-2015-022, (2015), http://cds.cern.ch/record/2037697.

[67] ATLAS Collaboration, *Commissioning of the ATLAS b-tagging algorithms using* $t\bar{t}$ *events in early Run-2 data*, ATL-PHYS-PUB-2015-039, (2015), http://cds.cern.ch/record/2047871.

[68] ATLAS Collaboration, *b-tagging calibration plots using dileptonic* $t\bar{t}$ *events produced in pp collisions at* $\sqrt{s} = 13$ *TeV and a combinatorial likelihood approach*, http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2016-002/.

[69] ATLAS Collaboration, *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at* $\sqrt{s} = 13$ *TeV*, ATL-PHYS-PUB-2015-023, (2015), https://cds.cern.ch/record/2037700.

[70] ATLAS Collaboration, *Luminosity determination in pp collisions at* $\sqrt{s} = 8$ *TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **76** (2016) 653, arXiv:1608.03953 [hep-ex].

[71] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at* $\sqrt{s} = 13$ TeV, ATL-PHYS-PUB-2016-015, (2016), http://cds.cern.ch/record/2203514.

[72] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Rev. D **96**

(2017) 072002, arXiv:1703.09665 [hep-ex].

[73] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, ATL-COM-PHYS-2016-407, (2016), https://cds.cern.ch/record/2149445.

[74] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, J. High Energy Phys. **11** (2004) 040, arXiv:hep-ph/0409146.

[75] S. Frixione et al., *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, J. High Energy Phys. **11** (2007) 070, arXiv:0709.2092 [hep-ph].

[76] S. Alioli et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, J. High Energy Phys. **06** (2010) 043, arXiv:1002.2581 [hep-ph].

[77] T. Melia et al., W^+W^-*, WZ and ZZ production in the POWHEG BOX*, J. High Energy Phys. **11** (2011) 078, arXiv:1107.5051 [hep-ph].

[78] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[79] T. Sjöstrand et al., *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv:1410.3012 [hep-ph].

[80] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, J. High Energy Phys. **07** (2002) 012, arXiv:hep-ph/0201195.

[81] T. Gehrmann et al., W^+W^- *Production at Hadron Colliders in Next to Next to Leading Order QCD*, Phys. Rev. Lett. **113** (2014) 212001, arXiv:1408.5243 [hep-ph].

[82] E. Bagnaschi et al., *Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM*, J. High Energy Phys. **02** (2012) 088, arXiv:1111.2854 [hep-ph].

[83] C. Anastasiou et al., *Higgs Boson Gluon-Fusion Production in QCD at Three Loops*, Phys. Rev. Lett. **114** (2015) 212001, arXiv:1503.06056 [hep-ph].

[84] F. Cascioli et al., *Precise Higgs-background predictions: merging NLO QCD and squared quark-loop corrections to four-lepton + 0,1 jet production*, J. High Energy Phys. **01** (2014) 046, arXiv:1309.0500 [hep-ph].

[85] F. Caola et al., *QCD corrections to* W^+W^- *production through gluon fusion*, Phys. Lett. B **754** (2016) 275, arXiv:1511.08617 [hep-ph].

[86] J. M. Campbell et al., *Top-Pair Production and Decay at NLO Matched with Parton Showers*, J. High Energy Phys. **04** (2015) 114, arXiv:1412.1828 [hep-ph].

[87] E. Re, *Single-top* Wt*-channel production matched with parton showers using the POWHEG method*, Eur. Phys. J. C **71** (2011) 1547, arXiv:1009.2450 [hep-ph].

[88] T. Sjöstrand et al., *PYTHIA 6.4 Physics and Manual*, J. High Energy Phys. **05** (2006) 026, arXiv:hep-ph/0603175.

[89] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, Comput. Phys. Commun. **185** (2014) 2930, arXiv:1112.5675 [hep-ph].

[90] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-*, Phys. Rev. D **82** (2010) 054018, arXiv:1005.4451 [hep-ph].

[91] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, J. High Energy Phys. **07** (2014) 079, arXiv:1405.0301 [hep-ph].

[92] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244, arXiv:1207.1303 [hep-ph].

[93] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A Precision tool for QED corrections in Z and W decays*, Eur. Phys. J. C **45** (2006) 97, arXiv:hep-ph/0506026.

[94] C. Anastasiou et al., *High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO*, Phys. Rev. D **69** (2004) 094008, arXiv:hep-ph/0312266.

[95] M. Grazzini et al., *$W^\pm Z$ production at hadron colliders in NNLO QCD*, Phys. Lett. B **761** (2016) 179, arXiv:1604.08576 [hep-ph].

[96] M. Grazzini et al., *ZZ production at the LHC: fiducial cross sections and distributions in NNLO QCD*, Phys. Lett. B **750** (2015) 407, arXiv:1507.06257 [hep-ph].

[97] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Methods **462** (2001) 152, BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines.

[98] ATLAS Collaboration, *Estimation of fake lepton background for top analyses using the Matrix Method with the 2015 dataset at $\sqrt{s}$=13 TeV with AnalysisTop-2.3.41*, ATL-COM-PHYS-2016-198, (2016), https://cds.cern.ch/record/2135116.

[99] ALEPH, DELPHI, L3 and OPAL Collaborations, the LEP Electroweak Working Group, *A Combination of Preliminary Electroweak Measurements and Constraints on the Standard Model*, arXiv:hep-ex/0612034.

[100] ATLAS Collaboration, *Measurement of total and differential W^+W^- production cross sections in proton-proton collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector and limits on anomalous triple-gauge-boson couplings*, J. High Energy Phys. **09** (2016) 029, arXiv:1603.01702 [hep-ex].

[101] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016)

023001, arXiv:1510.03865 [hep-ph].

[102] I. W. Stewart and F. J. Tackmann, *Theory Uncertainties for Higgs and Other Searches Using Jet Bins*, Phys. Rev. D **85** (2012) 034011, arXiv:1107.2117 [hep-ph].

[103] F. James and M. Roos, *Minuit: a system for function minimization and analysis of the parameter errors and correlations*, Comput. Phys. Commun. **10** (1975) 343.

[104] ATLAS Collaboration, *Measurement of the W^+W^- production cross section in pp collisions at a centre-of-mass energy of $\sqrt{s}$ = 13 TeV with the ATLAS experiment*, Phys. Lett. B **773** (2017) 354, arXiv:1702.04519 [hep-ex].

[105] T. Gehrmann et al., *W^+W^- Production at Hadron Colliders in Next to Next to Leading Order QCD*, Phys. Rev. Lett. **113** (2014) 212001, arXiv:1408.5243 [hep-ph].

[106] M. Grazzini et al., *W^+W^- production at the LHC: fiducial cross sections and distributions in NNLO QCD*, J. High Energy Phys. **08** (2016) 140, arXiv:1605.02716 [hep-ph].

[107] F. Caola et al., *QCD corrections to vector boson pair production in gluon fusion including interference effects with off-shell Higgs at the LHC*, J. High Energy Phys. **07** (2016) 087, arXiv:1605.04610 [hep-ph].

[108] F. Caola et al., *QCD corrections to W^+W^- production through gluon fusion*, Phys. Lett. B **754** (2016) 275, arXiv:1511.08617 [hep-ph].

[109] C. Anastasiou et al., *High precision determination of the gluon fusion Higgs boson cross-section at the LHC*, J. High Energy Phys. **05** (2016) 058, arXiv:1602.00695 [hep-ph].

[110] LHC Higgs Cross Section Working Group Collaboration, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, arXiv:1610.07922 [hep-ph].

[111] B. Biedermann et al., *Next-to-leading-order electroweak corrections to $pp \to W^+W^- \to 4$ leptons at the LHC*, J. High Energy Phys. **06** (2016) 065, arXiv:1605.03419 [hep-ph].

[112] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, arXiv:1007.1727 [physics.data-an], [Erratum: Eur. Phys. J. C **73** (2013) 2501].

[113] G. J. Feldman et al., *A Unified approach to the classical statistical analysis of small signals*, Phys. Rev. D **57** (1998) 3873, arXiv:physics/9711021 [physics.data-an].

[114] G. Corcella et al., *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, J. High Energy Phys. **01** (2001) 010, arXiv:hep-ph/0011363.

[115] J. M. Butterworth et al., *Multiparton interactions in photoproduction at HERA*, Z. Phys.

C **72** (1996) 637, arXiv:hep-ph/9601371.

[116] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, J. High Energy Phys. **06** (2002) 029, arXiv:hep-ph/0204244.